09/914665 .03 31 02

SUBSTITUTE FORM PTO-1390 U S DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER
10737-006001

U.S. APPLICATION NO. (If Known, see 37 CFR 1.5)
09/914665

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/EP00/02064	9 March 2000	10 March 1999

TITLE OF INVENTION
RETROVIRAL EXPRESSION VECTORS ON THE BASIS OF HERV-LONG TERMINAL REPEAT SEQUENCES

APPLICANT(S) FOR DO/EO/US
Christine Leib-Mösch, Ulrike Schön, Corinna Baust and Robert Michael Saller

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This is an express request to promptly begin national examination procedures (35 U.S.C. 371(f)).
4. ☐ The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. ☒ is attached hereto (required only if not communicated by the International Bureau).
   b. ☐ has been communicated by the International Bureau.
   c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☒ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. ☐ are attached hereto (required only if not communicated by the International Bureau).
   b. ☐ have been communicated by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☒ have not been made and will not be made.
8. ☐ An English language translation of amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☒ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11 to 16 below concern other documents or information included:

11. ☒ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. ☒ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.
14. ☐ A substitute specification.
15. ☐ A change of power of attorney and/or address letter.
16. ☐ Other items or information:
    ☐
    ☐
    ☐
    ☐
    ☐

CERTIFICATE OF MAILING BY EXPRESS MAIL Express Mail Label No EL696752561US

I hereby certify under 37 CFR §1 10 that this correspondence is being deposited with the United States Postal Service as Express Mail Post Office to Addressee with sufficient postage on the date indicated below and is addressed to the Commissioner for Patents, Washington, D C 20231

August 31, 2001 Samantha Bell Samantha Bell

Date of Deposit Signature Typed Name of Person Signing

[redacted]

518 Rec'd PCT/PTO 31 AUG 2001

U.S. APPLICATION NO. (IF KNOWN) 09/914665	INTERNATIONAL APPLICATION NO. PCT/EP00/02064	ATTORNEY'S DOCKET NUMBER 10737-006001

17. ☒ The following fees are submitted:

Basic National Fee (37 CFR 1.492(a)(1)- (5)):

	CALCULATIONS PTO USE ONLY	
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO **$1000**		
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO **$860**		
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO..... **$710**		
International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4)..... **$690**		
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4)..... **$100**		
ENTER APPROPRIATE BASIC FEE AMOUNT =	$860.00	
Surcharge of **$130** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).	$0.00	

Claims	Number Filed	Number Extra	Rate		
Total Claims	21 - 20 =	1	x **$18**	$18.00	
Independent Claims	2 - 3 =	0	x **$80**	$0.00	
MULTIPLE DEPENDENT CLAIMS(S) (if applicable)			+ **$270**	$0.00	
TOTAL OF ABOVE CALCULATIONS =				$878.00	
☐ Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.				$0.00	
SUBTOTAL =				$878.00	
Processing fee of **$130** for furnishing the English Translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(f))				$0.00	
TOTAL NATIONAL FEE =				$878.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property +				$0.00	
TOTAL FEES ENCLOSED =				$878.00	
				Amount to be refunded:	$
				Charged:	$

a. ☒ A check in the amount of $878.00 to cover the above fees is enclosed.
b. ☐ Please charge my Deposit Account No. 06-1050 in the amount of $0.00 to cover the above fees. A duplicate copy of this sheet is enclosed.
c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 06-1050. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Y. Rocky Tsao
FISH & RICHARDSON P.C.
225 Franklin Street
Boston, Massachusetts 02110-2804
(617) 542-5070 *phone*
(617) 542-8906 *facsimile*

Y. Rocky Tsao
SIGNATURE

Y. Rocky Tsao
NAME

34,053
REGISTRATION NUMBER

20309513.doc

Attorney's Docket No.: 10737-006001 / P13419-DrB/la

09/914665

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Christine Leib-Mösch et al.
Serial No. : Unassigned
Filed : Herewith
Title : RETROVIRAL EXPRESSION VECTORS ON THE BASIS OF HERV-LONG TERMINAL REPEAT SEQUENCES

BOX PCT
Commissioner for Patents
Washington, D.C. 20231

PRELIMINARY AMENDMENT

Applicants submit herewith a Sequence Listing in computer readable form as required by 37 CFR §1.824. In addition, applicants submit a substitute Sequence Listing as required under 37 CFR §1.823(a) and a statement under 37 CFR §1.821(f).

Applicants respectfully requests entry of the paper copy and computer readable copy of the Sequence Listing filed herewith for the instant application. Furthermore, applicants request entry of the following amendments.

In the specification:

Replace the original Sequence Listing with the substitute Sequence Listing filed herewith.

In the claims:

Amend claims 3-7, 9-12, 14, 15, and 20 as follows:

--3. (Amended) Vector according to claim 1 wherein the whole LTR region, the U3 region, or the R and U3 regions are derived from a human endogenous retroviral nucleotide sequence.--

CERTIFICATE OF MAILING BY EXPRESS MAIL

Express Mail Label No EL696752561US

I hereby certify under 37 CFR §1 10 that this correspondence is being deposited with the United States Postal Service as Express Mail Post Office to Addressee with sufficient postage on the date indicated below and is addressed to the Commissioner for Patents, Washington, D C 20231

August 31, 2001
Date of Deposit

Samantha Bell
Signature

Samantha Bell
Typed or Printed Name of Person Signing Certificate

--4. (Amended) Vector according to claim 1 wherein said nucleotide sequences encoding one or more proteins or elements of therapeutic and cytokin peptides are selected from one or more the group consisting of marker genes, therapeutic genes, antiviral genes, anti-tumor genes, and cytokin genes.--

--5. (Amended) Vector according to claim 1 wherein said cell-specifically controllable promoter region is derived from the LTR region of a cell-specifically expressed endogenous human retroviral nucleotide sequence.--

--6. (Amended) Vector according to claim 1 wherein said human endogenous retroviral cell-specifically controllable promoter sequences are selected from one or more promoter sequences of HERV families of the group consisting of HERV-K, HERV-H, HERV-E, HERV-L, HERV-T, HERV-R, HERV-I, HERV-P, ERV9, HERV-W.--

--7. (Amended) Vector according to claim 1 wherein said promoter region besides the TATA box additionally comprises recognition and binding sites for regulatory proteins.--

--9. (Amended) Vector according to claim 1 wherein said vector is a promoter conversion vector comprising a 5' LTR portion having the structure U3-R-U5, one or more sequences selected from coding and non-coding sequences, and a 3' LTR portion comprising a U3 region which is partially or completely deleted wherein the deleted U3 portion is replaced by a cell-specifically controllable promoter region from a HERV LTR sequence, followed by the R-U5 region.--

--10. (Amended) The mRNA or RNA of a retroviral expression vector according to claim 1.--

--11. (Amended) Prokaryotic cell or eukaryotic cell containing a retroviral expression vector according to claim 1.--

--12. (Amended) Eukaryotic cell containing a retroviral expression vector according to claim 1 in an integrated form.--

--14. (Amended) Use of an expression vector according to claim 1 for the expression of foreign genes in gene therapy.--

--15. (Amended) Virion containing a retroviral expression vector RNA obtained by transcription of an expression vector DNA according to claim 1.--

--20. (Amended) Retroviral vector system comprising a retroviral expression vector according to claim 1 and a packaging cell line comprising at least one retroviral or recombinant retroviral construct encoding for the packaging proteins of the retroviral expression vector.--

REMARKS

Amendments have been made to the specification to replace the original Sequence Listing from the related PCT application with an amended Sequence Listing wherein the general information (i.e., inventors, priority data, and attorney docket number) has been amended to reflect the information for the instant application.

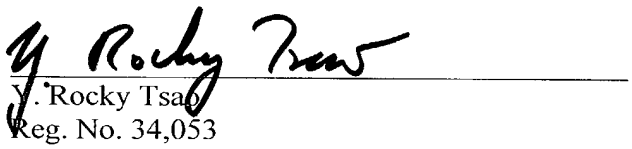
Amendments to the claims remove multiple dependency while conserving the claimed subject matter. No new matter has been introduced. Claims 1-21 are now pending. Applicants submit that all of the claims are now in condition for examination, which action is requested.

Attached is a marked-up version of the changes being made by the current amendment.

Please apply any charges or credits to Deposit Account No. 06-1050.

Respectfully submitted,

Date: 8-31-01



Y. Rocky Tsao
Reg. No. 34,053

Fish & Richardson P.C.
225 Franklin Street
Boston, Massachusetts 02110-2804
Telephone: (617) 542-5070
Facsimile: (617) 542-8906

Version with markings to show changes made

In the claims:

Claims 3-7, 9-12, 14, 15, and 20 have been amended as follows:

3. (Amended) Vector according to claim 1 [or claim 2] wherein the whole LTR region, the U3 region, or the R and U3 regions are derived from a human endogenous retroviral nucleotide sequence.

4. (Amended) Vector according to [one or more of the preceding claims] claim 1 wherein said nucleotide sequences encoding one or more proteins or elements of therapeutic and cytokin peptides are selected from one or more the group consisting of marker genes, therapeutic genes, antiviral genes, anti-tumor genes, and cytokin genes.

5. (Amended) Vector according to [one or more of the preceding claims] claim 1 wherein said cell-specifically controllable promoter region is derived from the LTR region of a cell-specifically expressed endogenous human retroviral nucleotide sequence.

6. (Amended) Vector according to [one or more of the preceding claims] claim 1 wherein said human endogenous retroviral cell-specifically controllable promoter sequences are selected from one or more promoter sequences of HERV families of the group consisting of HERV-K, HERV-H, HERV-E, HERV-L, HERV-T, HERV-R, HERV-I, HERV-P, ERV9, HERV-W.

7. (Amended) Vector according to [one or more of the preceding claims] claim 1 wherein said promoter region besides the TATA box additionally comprises recognition and binding sites for regulatory proteins.

9. (Amended) Vector according to [one or more of the preceding claims] claim 1 wherein said vector is a promoter conversion vector comprising a 5' LTR portion having the

structure U3-R-U5, one or more sequences selected from coding and non-coding sequences, and a 3' LTR portion comprising a U3 region which is partially or completely deleted wherein the deleted U3 portion is replaced by a cell-specifically controllable promoter region from a HERV LTR sequence, followed by the R-U5 region.

10. (Amended) The mRNA or RNA of a retroviral expression vector according to [one or more of the preceding claims] claim 1.

11. (Amended) Prokaryotic cell or eukaryotic cell containing a retroviral expression vector according to [one or more of the preceding claims] claim 1.

12. (Amended) Eukaryotic cell containing a retroviral expression vector according to [one or more of the preceding claims] claim 1 in an integrated form.

14. (Amended) Use of an expression vector according to [one or more of the preceding claims] claim 1 for the expression of foreign genes in gene therapy.

15. (Amended) Virion containing a retroviral expression vector RNA obtained by transcription of an expression vector DNA according to [one or more of the preceding claims] claim 1.

20. (Amended) Retroviral vector system comprising a retroviral expression vector according to [one or more of the preceding claims] claim 1 and a packaging cell line comprising at least one retroviral or recombinant retroviral construct encoding for the packaging proteins of the retroviral expression vector.

Attorney's Docket No.: 10737-006001 / P13419-DrB/la

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Christine Leib-Mösch et al.
Serial No. : Unassigned
Filed : Herewith
Title : RETROVIRAL EXPRESSION VECTORS ON THE BASIS OF HERV-LONG TERMINAL REPEAT SEQUENCES

BOX PCT
Commissioner for Patents
Washington, D.C. 20231

VERIFIED STATEMENT UNDER 37 CFR §1.821(f)

I, Jennifer H. Payne, declare that I personally prepared the paper and the computer-readable copy of the Sequence Listing filed herewith for the above-identified application and that the content of both is the same.

I further declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of The United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Date: August 31, 2001

Jennifer H. Payne

Fish & Richardson P.C.
225 Franklin Street
Boston, Massachusetts 02110-2804
(617) 542-5070 telephone
(617) 542-8906 facsimile

20309020.doc

CERTIFICATE OF MAILING BY EXPRESS MAIL

Express Mail Label No. EL 696 752 561 US

I hereby certify under 37 CFR §1.10 that this correspondence is being deposited with the United States Postal Service as Express Mail Post Office to Addressee with sufficient postage on the date indicated below and is addressed to the Commissioner for Patents, Washington, D C. 20231.

August 31, 2001
Date of Deposit

Samantha Bell
Signature

Samantha Bell
Typed or Printed Name of Person Signing Certificate

SEQUENCE LISTING

<110> Christine Leib-Mösch
Ulrike Schön
Corinna Baust

<120> RETROVIRAL EXPRESSION VECTORS ON THE BASIS OF HERV-LONG TERMINAL REPEAT SEQUENCES

<130> 10737-006001

<150> PCT/EP00/02064
<151> 2000-03-09

<150> DE 199 10 650.9
<151> 1999-03-10

<160> 47

<170> PatentIn Ver. 2.1

<210> 1
<211> 375
<212> DNA
<213> Human endogenous retrovirus

<400> 1
tgtcaggcct ctgagcccaa gctaagccat catatcccct gtgacctgca cgtatacatc 60
cagatagcct gaagcaactg taaaaatatc cttaactgat gacattccaa tattgtgatt 120
tgtttctgcc ctaccctgac tgatcaatgt gctttgtaat ctcccccacc cttcagaagg 180
ctctttgtaa tcctccccac ccttgagaat ggacttggtg agatccaccc cctgcctgca 240
aagcattgcc cctaactcca ccgcctgtcc caaaacctat gagaactaat gataatccca 300
ccacactttg ctgactctct ttccagactc agcccggctg cacccaggtg aaataaacag 360
ccatgttgct cacat 375

<210> 2
<211> 400
<212> DNA
<213> Human endogenous retrovirus

<400> 2
tgtcaggcct ctgagcccaa gccaggccat cgcatcccct gtgacttgca cgtatacatc 60
cagatggcct aaagtaactg aagatccaca aaagaagtaa aaacagcctt aactgatgac 120
attccaacat tgtgatttgt tcctgcccca ccctaactga taaatgtact ttgtaatctc 180
ccccaccctt aagaaggtcc tttgtaattc tccccaccct tgagagtgta ctttgtgaga 240
tccacacctg cccaccagag aacaaacccc ctttgactgt aattttccat taccttccct 300
aatcctataa aacggcccca ccccatctcc ctttgctgac tctcttttcg gactcagccc 360
gcctgcaccc aggtgaaata aacagccatg ttgctcacat 400

<210> 3
<211> 361
<212> DNA
<213> Human endogenous retrovirus

<400> 3

tgtcaggcct ctgagcccaa gctaagccat catatcccct gcgacctgca catatacatc 60
cagatggcct gaagtaactg aagaatcaca aaagaagtga aaatggcctg ttcctgcctt 120
aactgatgac attaccttgt gaaattcctt ctcctggctc atcctggctc aaaagctccc 180
gcactgagca ccttgtgacc cctgcctctg cccgccagag agcaaccccc tcttgactgt 240
aattttcctt tacctaccta aatcctataa aatggcccca ctcctatctc ccttcgctga 300
ctctcttttc ggactcagcc cgcctgtacc caggtgaaat aaacagccat gttgctcaca 360
t 361

<210> 4
<211> 372
<212> DNA
<213> Human endogenous retrovirus

<400> 4
tgtcaggcct ctgagcccaa gctaagccat catatcccct gtgacctgca cgtacacatc 60
cagatggccg gttcctgcct taactgatga cattccacca cgaaagaagt gaaaatgacc 120
tgttcgtgcc ttaactgatg acattgtctt gtgaaattcc tcctcctggc tcatcctggc 180
tcaaaagctc cccgactgag tacattgtga cccccactcc tgcccgccag agaacagccc 240
cctttgactg taattttcct ttatctaccc aaatcctata aaacagcccc acctttatct 300
cccttggctg actctctttt cggactcagc ccgcctgcac ccaggtgaaa taaacagcca 360
tgttgctcac at 372

<210> 5
<211> 349
<212> DNA
<213> Human endogenous retrovirus

<400> 5
tgtcaggcct ctgagcccaa gccaagccat cgcatcccct gtgacttgca cgtgtatgcc 60
cagatggcct gaagtaactg aagaatcaca aaagaagtga aaaggccctg ccccacctta 120
actgagtgat taaccccatg aatttccttc ccctggctca gaagctcccc cactgagcac 180
cttgtgaccc ctgcccctgc ccaccagaga acaaccccct ttgactgtaa ttttccatta 240
ctttcccaaa tcctataaaa cggccccacc cctatctccc ttcgctgact ctcttttcgg 300
actcagcccg cctgcaccca ggtgaaataa acagccatgt tgctcacat 349

<210> 6
<211> 324
<212> DNA
<213> Human endogenous retrovirus

<400> 6
tgtcaggcct ctgagcccaa gcctgcacgt atacatccag atgaagcaag tgaagaatca 60
caaaagaagt gaaaatggcc ggttcctgcc ttaactgatg acattacctt gtgaaattcc 120
ttctcctggc tcagaagctc ccccactgag caccttgtga cccccactcc tccccgccac 180
agaacaaccc cctttgactg taattttcca ctgcccgccc aaaccctata aaacggtccc 240
accccatctc ccttccctga ctctcttttc ttcggactca gcccgcctgc acccaggtga 300
aataaacagc catgttgctc acat 324

<210> 7
<211> 393
<212> DNA
<213> Human endogenous retrovirus

<400> 7
tgtcaggcct ctgagcccaa gccaagccat cgcatcccct gtgacttgca cgtatacgcc 60
cagatggcct gaagtaactg aagaatcaca aaagaactga aaaggccctg ccccgcctta 120
actgatgaca ttccaccatg gtgatttgtt cttgccccac cttaactgag tgattaaccc 180
tgtgaatttg cttctcctgg ctcagaagct cccccactga gcaccttgtg acccccgccc 240
ctgcccacca gagaacaacc ccctttgact gtaattttcc attaccttcc caaatcctat 300
aaaacggccc cacccctatc tcccttcgct gactctcttt tcggactcag cccgcctgcc 360
cccaggtgaa ataaacagcc atgttgctca cat 393

<210> 8
<211> 393
<212> DNA
<213> Human endogenous retrovirus

<400> 8
tgtcaggcct ctgagcccaa gccaagccat cgcatcccct gtgacttgca cgtatacgcc 60
cagatggcct gaagtaactg aagaatcaca aaagaactga aaaggccctg ccccgcctta 120
actgatgaca ttccaccatg gtgatttgtt cttgccccac cttaactgag tgattaaccc 180
tgtgaatttg cttctcctgg ctcagaagct cccccactga gcaccttgtg acccccgccc 240
ctgcccacca gagaacagac ccctttgact gtaattttcc attaccttcc caaatcctat 300
aaaacggccc cacccctatc tcccttcgct gactctcttt tcggactcag cccgcctgcc 360
cccaggtgaa ataaacagcc atgttgctca cat 393

<210> 9
<211> 388
<212> DNA
<213> Human endogenous retrovirus

<400> 9
tgtcaggcct ctgagcccaa gctaagccat catatcccca gggacctgca cgtatacatc 60
cagatggcct gaagcaactg aagatccaca aaggaagtga aaatagcctt aactgatgac 120
attccaccat tgtgatttgt ttctgcccca tcctaactga tcaatgtact ttgtaatctc 180
tcccaccctt aagaaggttc tttgtaattc tccccaccct tgagagtgta ctttgtgaga 240
tccacccccc gccggcaaaa cattgctcct aacccaaccg cctaccccaa acctgtaaga 300
actaatgata atccaccacc ctttgctgac tcttttcaga atcagcccgc ctgcacccag 360
gtgaaataaa cagccatgtt gctcacat 388

<210> 10
<211> 314
<212> DNA
<213> Human endogenous retrovirus

<400> 10
tgtcaggcct ctgagcccaa gctaagccat caaatcccct gtgacctgca cgtgtacatc 60
cagatgacct gaagcaactg aagatccaca aaagaagtga aagtagcctt aactgatgac 120
attccaccat tgtgatttgt tcctgcccca cgctaactga taccatatat tcttcccccg 180
cccttgagaa tgtactttgt acacctatcc caaacctata agaactaatg ataatcctac 240
cacccttgc tgactctctt tttggactca gcccgcctgc acccaggtga aataaacagc 300
catgttgctc acat 314

<210> 11
<211> 309
<212> DNA

<213> Human endogenous retrovirus

<400> 11
tgtcaggcct ctgagcccaa gctaagccat catatcccgt gacctgcata tacatccaga 60
tggcctgaag caactgaaga tccacaaagg aagtgaaaat agccttaact gatgacattc 120
caccattgtg atttgttcct gccccacgct aactgatacc atatattctt cccccgccct 180
tgagaatgta ctttgtacac ctatcccaaa cctataagaa ctaatgataa tccaccaccc 240
tttgctgact ctcttttgg actcagcccg cctgcaccca ggtgaaataa acagccatgt 300
tgctcacat 309

<210> 12
<211> 314
<212> DNA
<213> Human endogenous retrovirus

<400> 12
tgtcaggcct ctgagcccaa gctaagccat caaatcccct gtgacctaca cgtgtacatc 60
cagatgacct gaagcaactg aagatccaca aaagaagtga aagtagcctt aactgatgac 120
attccaccat tgtgatttgt tcctgcccta cgctagctga taccatatat tcttcccccg 180
cccttgagaa tgtactttgt acacctatcc caaacctata agaactaatg ataatcctac 240
cacccttgc tgactctctt tttggactca gcccgcctgc acccaggtga aataaacagc 300
catgttgctc acat 314

<210> 13
<211> 341
<212> DNA
<213> Human endogenous retrovirus

<400> 13
tgttgagatg ggggactgag agacaggact agctggattt cctaggccga ctaagaatcc 60
ctaagcctag ctgggaaggt gaccgcatcc acctttaaac acggggctcg caacttagct 120
cacacccaac caatcaggta gtaaagaggg ctcactaaaa tgctaattag gcaaagacag 180
gaggtaaaga aatagccaat catctattgc ctgagagcac agcaggaggg acaatgatcg 240
ggatataaac ccaagtcttc gagccggcaa tggctacctt ctttgggtcc cctcccttg 300
tatgggagct ctgttttcac tctattaaat cttgcaactg c 341

<210> 14
<211> 341
<212> DNA
<213> Human endogenous retrovirus

<400> 14
tgttgagatg ggggactgag agacaggact agctggattt cctaggccga ctaagaattc 60
ctaagcctag ctgggaaggt gaccgcatcc atctttaaac atggggcttg caacttaact 120
catatctgac caatcaggta gtaaagagag ctcactaaaa tgctaattag gctaaaacag 180
gaggcaaaga agtagccaat catctgttgc ctgacagcac agcaggaggg acaatgatcg 240
ggatataaac ccaggcattc gagccagcta cagctaccct ctttgggtcc cctcccttg 300
tatgggagct ctgtcttcac tctattaaat cttgcaactg c 341

<210> 15
<211> 322
<212> DNA
<213> Human endogenous retrovirus

<400> 15
tgttgagatg ggggactgag agacaggact acctggattt cctaggccga ctaagaatcc 60
ctaagcctag ctgggaaggt gaccacatcc acctttaaac acagggcttg caacttagct 120
cacacttgac cagtcaggta gtaaagagag ctcactaaaa tgctaattag gctaaaacag 180
gaggtaaaga aatagacaat catctatcac ctgagagcac agtgggaggg acaatgatcg 240
gcatataaac ccaggcattc gagccagcaa cagcaacccc ctttgggagc tctgttttca 300
ctctattaaa tcttgcaact gc 322

<210> 16
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 16
tgttgagatg ggggactgag agacaggact agctggattt cctaggccaa ctaagaatcc 60
ctaagcctag ctgggaaggt gactacaccc acctttaaac atggggcttg caacttagct 120
cacacccaac caatcaggta gtaaagagag cttgctaaaa tgctaattag gcaaaaacag 180
gaggtaaaga aatagccagt catctatcgc ctgacagcac aaggggcggg acaatgatca 240
ggatataaac tcaggcattc aagccagcaa tggctaccca ctttgggtcc cctcccattt 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 17
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 17
tgttgagatg ggggactgag agacaggact agctggattc cctaggccga ctaagaatcc 60
ctaagcctag ctgggaaggt gaccacatcc acctttaaac acggggcttg caacttagct 120
catacccaac aaatcaggta gtaaagagag ctcactaaaa tactgattag gcgaaaacag 180
gaggtaagga aacagccagt catctatcgc ctgacagcac aaggggcggg acaatgatca 240
ggatataaac tcaggcattc aagccagcaa tggctaccca ctttgggtcc cctcccattt 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 18
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 18
tgttgagatg ggggactgag agacaggact agttggattt cctaggctgg ctaagaatcc 60
ctaagcctag ctgggaaatt gaccacgtcc acctttaaac acggggcttg caatttagct 120
cacacccgac caatcaggta gtaaagggag ctcactaaaa tgctaattag ggaaaaacag 180
gaggtaaaga agtagccaat catctatcgc ctgagagcac aacaggaggg acaatgatca 240
ggatataaac ccaggcattc aagccagcgg tggctaccct ctttgggtcc cctccctttg 300
tatgggagcc ctgttttcac tctattaaat cttgcaactg caa 343

<210> 19
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 19
tgttgagatg ggggactgag agacaggact agttggattt cctaggccgg ctaagaatcc 60
ctaagcctag ctgggaaatt gaccacgtcc acctttaaac acggggcttg caatttagct 120
cacacccgac caatcaggta gtaaagggag ctcactaaaa tgctaattag ggaaaaacag 180
gaggtaaaga agtagccaat catctatcgc ctgagagcac aacaggaggg acaatgatca 240
ggatataaac ccaggcattc aagccagcgg tggctaccct ctttgggtcc cctccctttg 300
tatggaagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 20
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 20
tgttgagatg ggggactgag agacaggact agctggattt cctaggccaa ctaagaatcc 60
ctaagcctag ctgggaaggt gactacaccc acctttaacc actaggcttg caacttagct 120
cacacccgac caatcaggta gtaaagagag cttgctaaaa tgctaattag gcaaaaacag 180
gaggtagaga aatagccaat catctatcgc ctgagagcac agcaggaggg acaatgatcc 240
ggatataaac ccaagcattc gagccagcaa tggctaccct ctttgtgtcc cctccctttg 300
tatgggagct ctattttcac tctattaaat cttgcaactg caa 343

<210> 21
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 21
tgttgagatg ggggactgag agacaggact agctggattt cctaggctga ctaagaatcc 60
ctaagcctag ctgggaaggt gaccgcatcc atctttaaac atggggcttg caacttaact 120
catatctgac caatcaggta gtaaagagag cttgctaaaa tgctaattag gcaaaaacag 180
gaggtaaaga aatagccagt catctatcgc ctgacagcac aaggggcggg acaatgatca 240
ggatataaac tcaggcattc aagccagcaa tggctaccca ctttgggtcc cctcccattt 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 22
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 22
tgttgagatg ggggactgag agacaggact agctggattt cctaggctga ctaagaatcc 60
ctaagcctag ctgggaaggt gactacaccc acctttaacc actaggcttg caacttagct 120
cacacccgac caatcaggta gtaaagagag cttgctaaaa tgctaattag gcaaaaacag 180
gaggtaaaga aatagccagt catctatcgc ctgacagcac aaggggcggg acaatgatca 240
ggatataaac tcaggcattc aagccagcaa tggctaccca ctttgggtcc cctcccattt 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 23
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 23

tgttgagatg ggggactgag agacaggact agttggattt cctaggctgg ctaagaatcc 60
ctaagcctag ctgggaaatt gaccacgtcc acctttaaac acggggcttg caatttagct 120
cacacccgac caatcaggta gtaaagggag ctcactaaaa tgctaattag ggaaaaacag 180
gaggtaaaga agtagccaat catctatcgc ctgagagcac aacaggaggg acaatgatca 240
ggatataaac ccaggcattc aagccagcgg tggctaccct ctttgggtcc cctccctttg 300
tatggaagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 24
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 24
tgttgagatg ggggactgag agacaggact acctggattt cctaggccaa ctaagaatct 60
ctaagcctag ctgggaaggt gaccacatcc acctttaaac acagggcttg caacttagct 120
cacacccgac ccatcaggta agaaagagag cccgctaaaa tgctaattag gcaaaaacag 180
gaggtaaaga aatagtcaat catctattgc ctgagagcac agcgggaggg acaatgatca 240
ggatataaac ccaggcattc gagccggcaa cgactaccct ctttgggtcc cctccctttg 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 25
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 25
tgttgagatg ggggactgag agacaggact agctggattt cctaggccaa ctaagaatcc 60
ctaagcctag ctgggaaggt gactacaccc acctttaacc actaggcttg caacttagct 120
cacacccgac caatcaggta gtaaagagag cttgctaaaa tgctaattag gcaaaaacag 180
gaggtaaaga aatagccagt catctatcgc ctgacagcac aaggggcggg acaatgatca 240
ggatataaac tcaggcattc aagccagcaa tggctaccca ctttgggtcc cctcccattt 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 26
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 26
tgttgagatg ggggactgag aaacaggact agcaggattt cctaggccga ttaagaatcc 60
ctaagcctag atgggaagtt gaccacatcc acctttaaac acggggcttg caactcagct 120
cacacccgac ccatcaggta agaaagagag cccgctaaaa tgctaattag gcaaaaacag 180
gaggtaaaga aatagccaat catctattgc ctgagagcac agcgggaggg acaatgatca 240
ggatataaac ccaggcattc gagccggcaa cgactaccct ctttgggtcc cctccctttg 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 27
<211> 619
<212> DNA
<213> Human endogenous retrovirus

<400> 27
gcgaccggtg gatcccgggc ccgcggtacc gtcgactgca gaattcatgg agcatacaat 60

cgggttttat accgagacat tccattgccc agggacaggc aggagacaga tgccttcctc 120
ttgtctcaac tgcaagaggc attccttcct cttatactaa tcctcctcag cacagaccct 180
ttacgggtgt cgggctgggg gacggtcagg tctttccctt cccacgaggc catatttcag 240
actatcacat ggggagaaac cttggacaat acctggcttt cctaggcaga ggtccctgcg 300
gccttccgca gtttttgtgt cctgggtact tgagattagg gagtggtgat gactcttaag 360
gagcatgctg ccttcaagca tctgtttaac aaagcacatc ctgcaccgcc cttaatccat 420
tcaaccctga gttgacacag cacacgtttc agagagcacg gggttggggg taaggtcata 480
gattaacaga atctcaaggc agaagaattt ttcttaacac ataacaaaat ggagtctccc 540
atgtctactt ctttctacac agacacagta acaatctgat ctctcttgct tttccccaca 600
tttccccctt ttcttttcg 619

<210> 28
<211> 620
<212> DNA
<213> Human endogenous retrovirus

<400> 28
gcgaccggtg gatcccgggc ccgcggtacc gtcgactgca gaattcatgg agcatacaat 60
cgggttttat accgagacat tccattgccc agggacaggc aggagacaga tgccttcctc 120
ttgtctcaac tgcaagaggc attccttcct cttatactaa tcctcctcag cacagaccct 180
ttacgggtgt cgggctgggg gacggtcagg tctttccctt cccacgaggc catatttcag 240
actatcacat ggggagaaac cttggacaat acctggcttt cctaggcaga ggtccctgcg 300
gccttccgca gtttttgtgt cctgggtact tgagattagg gagtggtgat gactcttaag 360
gagcatgctg ccttcaagca tctgtttaac aaagcacatc ctgcaccgcc cttaatccat 420
tcaaccctga gttgacacag cacacgtttc agagagcacg gggttggggg taaggtcata 480
gattaacaga atctcaaggc agaagaattt ttcttaacac ataacaaaat ggagtctccc 540
atgtctactt ctttctacac agacacagta acaatctgat ctctcttgct tttccccaca 600
tttccccctt ttcttttcga 620

<210> 29
<211> 624
<212> DNA
<213> Human endogenous retrovirus

<400> 29
gcgaccggtg gatcccgggc ccgcggtacc gtcgactgca gaattcatgg agcatacaat 60
cgggttttat accgagacat tccattgccc agggacaggc aggagacaga tgccttcctc 120
ttgtctcaac tgcaagaggc attccttcct cttatactaa tcctcctcag cacagaccct 180
ttacgggtgt cgggctgggg gatggtcagg tctttccctt cccacgaggc catatttcag 240
actatcacat gggaagaaac cttggacaat acctggcttt cctaggcaga ggtccctgcg 300
gccttccgca gtttttgtgt cctgggtact tgagattagg gagtggtgat gactcttaag 360
gagcatgctg ccttcaagca tctgtttaac aaagcacatc ctgcactgcc cttaatccat 420
tcaaccctga gttgacacag cgcacgtttc agagagcacg gggttggggg taaggtcata 480
gattaacaga atctcaaggc agaagaattt ttcttaacac ataacaaaat ggagtctccc 540
atgtctactt ctttctacac agacacagta acaatctgat ctctcttgct tttccccaca 600
tttccccctt ttcttttcga caaa 624

<210> 30
<211> 646
<212> DNA
<213> Human endogenous retrovirus

<400> 30
gcgaccggtg gatcccgggc ccgcggtacc gtcgactgca gaattcatgg agcatacaat 60

[redacted]

cgggttttat accgagacat tccattgccc agggacaggc aggagacaga tgccttcctc 120
ttgtctcaac tgcaagaggc attccttcct cttatactaa tcctcctcag cacagaccct 180
ttacgggtgt cgggctgggg gacggtcagg tctttccctt cccacgaggc catatttcag 240
actatcacat ggggagaaac cttggacaat acctggcttt cctaggcaga ggtccctgcg 300
gccttccgca gtttttgtgt cctgggtact tgagattagg gagtggtgat gactcttaag 360
gagcatgctg ccttcaagca tctgtttaac aaagcacatc ctgcaccgcc cttaatccat 420
tcaaccctga gttgacacag cacacgtttc agagagcacg gggttggggg taaggtcata 480
gattaacaga atctcaaggc agaagaattt ttcttaacac ataacaaaat ggagtctccc 540
atgtctactt ctttctacac agacacagta acaatctgat ctctcttgct tttccccaca 600
tttccccctt ttcttttcga caaaaccgcc atctcgagat ctgagt 646

<210> 31
<211> 672
<212> DNA
<213> Human endogenous retrovirus

<400> 31
gtcccacctc cagccctaag gcggtttttc cctatctcag tagatggagc atacaatcgg 60
gttttatacc gagacattcc attgcccagg gacaggcagg agacagatgc cttcctcttg 120
tctcaactgc aagaggcatt ccttcctctt atactaatcc tcctcagcac agacccttta 180
cgggtgtcgg gctgggggac ggtcaggtct ttcccttccc acgaggccat atttcagact 240
atcacatggg gagaaacctt ggacaatacc tggctttcct aggcagaggt ccctgcggcc 300
ttccgcagtt tttgtgtcct gggtacttga gattagggag tggtgatgac tcttaaggag 360
catgctgcct tcaagcatct gtttaacaag gcacatcctg caccgccctt aatccattca 420
accctgagtt gacacagcac acgtttcaga gagcacgggg ttgggggtaa ggtcatagat 480
taacagaatc tcaaggcaga agaatttttc ttaacacata acaaaatgga gtctcccatg 540
tctacttctt tctacacaga cacagtaaca atctgatccc tcttgctttt ccccacattt 600
ccccctttc ttatccatca cactggcggc cgctcgagca tgcatctaga gggcccaatt 660
cgccctatag tg 672

<210> 32
<211> 593
<212> DNA
<213> Human endogenous retrovirus

<400> 32
agtagatgga gcatacaatc gggttttata ccgagacatt ccattgccca gggacaggca 60
ggagacagat gccttcctct tgtctcaact gcaagaggca ttccttcctc ttttactaat 120
cctcctcagc acagaccctt tacaggtgtc gggctggggg acggtcaggt ctttcccttc 180
ccacgaggcc atatttcaga ctatcacatg gggagaaacc ttggacaata cctggctttc 240
ctaggcagag gtccctgcgg ccttctgcag tttttgtgtc cctgggtact tgagattagg 300
gagtggtgat gactcttaag gagcatgctg ccttcaagca tctgtttaac aaagcacatc 360
ctgcaccgcc cttaatccat tcaaccctga gttgacacag cacatgtttc agagagcacg 420
gggttggggg taaggtcata gattaacaga atctcaaggc agaagaattt ttcttagcac 480
ataacaaaat ggagtctcct atgtctactt ctttctacac agacacagta acaatttgat 540
ctctcttgct tttccccaca tttccccctt ttcttttcga caaaaccgcc atc 593

<210> 33
<211> 943
<212> DNA
<213> Human endogenous retrovirus

<400> 33
tgtgggcgaa ggattaccca ggtgccgagg caagagactg aaggcacaaa ctgtttcagt 60

ataatataga aaatagctag aataagaata gttataataa aaattagata tacacatgat 120
catggacatt accaatcatt actacaaaca ttgttaatca ttagctttta atattactct 180
ttgttttatt actaatataa ccaaggaata accggtagca tacggtcagg tgctgaaggg 240
acattgtgag aagtgaccta gaaggcaaga ggtgagcctt ctgtcacgcc tgcataagga 300
cagcttgagg gctccttggt caagctgtaa caccagtgcc tgggaaggca ccgttactta 360
gcagaccatg aaagggagtc tccattcctt ggaggagtca gggaaacact atgctccacc 420
agcttcttgt gtatccagcc ctgcccacag tcatccagag gcataaaccc ctccctgtgg 480
tgctgtgctt caatggccat gcttcttgtc cactttcatg ttcctcctgt actcctggtt 540
cctctttgaa gttcgtagaa gataatggta gaagaaatag tgaaagtctt tgatctttct 600
tataagtgca tagaagaaaa cactgatgta tgcctgcctt ccctctctgc ttcagctacc 660
taaaaggaaa ggcccccttt cccatgatca catgacttgc ctgaccttat caatcacttg 720
gaggactcac cctccttacc ctgtcccttt gtcttgtatg caataaatat cagcacgccc 780
agccattcgg ggccactact ggtctccgca acttggtggt agtggtaccc tgggcccagc 840
tgttttctct ttatctcttt tgtcttgtgt ctttatttct tacaatctct catctctgca 900
catggggaga acaccggcaa agcccgtagg gctggacctt aca 943

<210> 34
<211> 389
<212> DNA
<213> Human endogenous retrovirus

<400> 34
aaacccctcc ctgtggtgct gtgcttcaat ggccatgctt cttgtccact ttcatgttcc 60
tcctgtactc ctggttcctc tttgaagttc gtagaagata atggtagaag aaatagtgaa 120
agtctttgat ctttcttata agtgcataga agaaaacact gatgtatgcc tgccttccct 180
ctctgcttca gctacctaaa aggaaaggcc ccctttccca tgatcacatg acttgcctga 240
ccttatcaat cacttggagg actcaccctc cttaccctgt ccctttgtct tgtatgcaat 300
aaatatcagc acgcccagcc attcggggcc actactggtc tccgcaactt ggtggtagtg 360
gtaccctggg cccagctgtt ttctcttta 389

<210> 35
<211> 858
<212> DNA
<213> Human endogenous retrovirus

<400> 35
tgtgggcgga agagtaccta ggtgccgagg caagagactg aaggcacaaa ctgtttcagt 60
ataataaaga aaatagaata agaatagtca taatacaaat tagatacagc gatgatcatg 120
aacaattatc catcattatt ataaacatta ttaatcatta gcttttaata ttactctgtt 180
gcattaataa tataacctag gaataaccgg caggtatagg gtcaggtgct gaagggacat 240
tgtgagaagt gaatagaagg caagagggga gccttctgtc atgcccgcat aagggccgct 300
tgagggcccc ttggtcaagc ggtaacgcca gtgtctggga aggcacccgt tactgagcag 360
accgggaaag ggagtctcct ttccttggag gagtcaggga acgctctgct ccaccagctt 420
cttgtgggag gctggatgtt acccaggcct gcctgcagtc atccggaggc ctgaacccct 480
ccctgtggtg cttcaatggt cacgttcctt gtccactttc atgctccttc cgtactcctg 540
gttcctcttt gaagttcgta gtagatagcg gtagaagaaa tagtgaaagt cttaaagtct 600
ttgatcttat aagttcatag aagaaaacgc tgatgcctgc cgccttctct ctctgcttca 660
gctacctaag agggaagggc ccgctgtcct gtgatcaggt gacttgcttc accttgtcaa 720
tcacttagaa gactgaccct ccttatcctg ccccttgtc ttgtatgcaa taaatatcag 780
cgagcccagc cgttcagggc cactaccggt ctccgtgtct ttgtggtagt ggtccccggg 840
cccagctgtt ttctcttt 858

<210> 36
<211> 386

<212> DNA
<213> Human endogenous retrovirus

<400> 36
gaacccctcc ctgtggtgct tcaatggtca cgttccttgt ccactttcat gctccttccg 60
tactcctggt tcctctttga agttcgtagt agatagcggt agaagaaata gtgaaagtct 120
taaagtcttt gatcttataa gttcatagaa gaaaacgctg atgcctgccg ccttctctct 180
ctgcttcagc tacctaagag ggaagggccc gctgtcctgt gatcaggtga cttgcttcac 240
cttgtcaatc acttagaaga ctgaccctcc ttatcctgcc cccttgtctt gtatgcaata 300
aatatcagcg agcccagccg ttcagggcca ctaccggtct ccgtgtcttt gtggtagtgg 360
tccccgggcc cagctgtttt ctcttt 386

<210> 37
<211> 844
<212> DNA
<213> Human endogenous retrovirus

<400> 37
tgtgggtgga ggattaccca ggtgccaagg caagagactg aaggcacaaa ctgtttcagt 60
ataataaaaa aaatagaata agaatagtca taatacaaat tagatataga gatgatcatg 120
gacaattagc aatcactatt aatctttagc ttttaatatt actctttgtt gcattactaa 180
tataacctag gaataaccgg tgggtatagg gtcaggtgct gaagggacat tgtgtgaagt 240
gacctggaag gcaagaggtg agccctctgt cacgcccaca taagggccgc ttgagggctc 300
cttggtcaag tggtaacgcc agtgtctggg aatgcacccg ttaattagca gaccgcgaaa 360
gggagtctcc tttccttgga agagttgggg aacactctgc tccaccagct tcttgtggaa 420
ggctggatat tatccaggcc tgcgcgcagt catccggagg cttaaaccct tccctgtggt 480
gctgtgcttc aatggtccca ctccttgtcc actttcatgc tcctcccgta ctcctggttc 540
ctctttgaag agcgcagtag atagcggtag aagaaatagt gaaagtctta aagtcttcga 600
tctttcttac aagtgcagag aagaaaacgc tgacatatgc tgccttccct ctctgcttcg 660
gctacctaaa agggaagggc cgcctatcct gtaatcacat gacttgcttc accttgtcaa 720
tcacttagaa gattcactct ccttaccctg ccccccttgtc ttgtatgcaa taaatatcag 780
tgaccccagc cgttcagggc cactactggt ctccgcgtct tgatggtagt ggtcaccccg 840
gccc 844

<210> 38
<211> 381
<212> DNA
<213> Human endogenous retrovirus

<400> 38
aaacccttcc ctgtggtgct gtgcttcaat ggtcccactc cttgtccact ttcatgctcc 60
tcccgtactc ctggttcctc tttgaagagc gcagtagata gcggtagaag aaatagtgaa 120
agtcttaaag tcttcgatct ttcttacaag tgcagagaag aaaacgctga catatgctgc 180
cttccctctc tgcttcggct acctaaaagg gaagggccgc ctatcctgta atcacatgac 240
ttgcttcacc ttgtcaatca cttagaagat tcactctcct taccctgccc ccttgtcttg 300
tatgcaataa atatcagtga ccccagccgt tcagggccac tactggtctc cgcgtcttga 360
tggtagtggt caccccggcc c 381

<210> 39
<211> 859
<212> DNA
<213> Human endogenous retrovirus

<400> 39

tgtgggtgga ggattaccca ggtgccaagg caagagactg aaggcacaaa ctgtttcagt 60
ataataaaaa aaatagaata agaatagtca taatacaaat tagatataga gatgatcatg 120
gacaattagc aatcactatt aatctttagc ttttaatatt actctttgtt gcattactaa 180
tataacctag gaataaccgg tgggtatagg gtcaggtgct gaagggacat tgtgagaagt 240
gacctggaag gcaagaggtg agccctctgt cacgcccaca taagggccgc ttgagggctc 300
cttggtcaag tggtaacgcc agtgtctggg aatgcacccg ttaattagca gaccgcgaaa 360
gggagtctcc tttccttgga agagttgggg aacactctgc tccaccagct tcttgtggaa 420
ggctggatat tatccaggcc tgcgcgcagt catccggagg cttaaaccct tccctgtggt 480
gctgtgcttc aatggtccca ctccttgtcc actttcatgc tcctcccgta ctcctggttc 540
ctctttgaag agcgcagtag atagcggtag aagaaatagt gaaagtctta aagtcttcga 600
tctttcttac aagtgcagag aagaaaacgc tgacatatgc tgccttccct ctctgcttcg 660
gctacctaaa agggaagggc cgcctatcct gtaatcacat gacttgcttc accttgtcaa 720
tcacttagaa gattcaccct ccttaccctg ccccccttgtc ttgtatgcaa taaatatcag 780
tgaccccagc cgttcagggc cactactggt ctccgcgtct tgatggtagt ggtcaccccg 840
gcccaggtgt tttttcttt 859

<210> 40
<211> 396
<212> DNA
<213> Human endogenous retrovirus

<400> 40
aaacccttcc ctgtggtgct gtgcttcaat ggtcccactc cttgtccact ttcatgctcc 60
tcccgtactc ctggttcctc tttgaagagc gcagtagata gcggtagaag aaatagtgaa 120
agtcttaaag tcttcgatct ttcttacaag tgcagagaag aaaacgctga catatgctgc 180
cttccctctc tgcttcggct acctaaaagg gaagggccgc ctatcctgta atcacatgac 240
ttgcttcacc ttgtcaatca cttagaagat tcaccctcct taccctgccc ccttgtcttg 300
tatgcaataa atatcagtga ccccagccgt tcagggccac tactggtctc cgcgtcttga 360
tggtagtggt caccccggcc caggtgtttt ttcttt 396

<210> 41
<211> 966
<212> DNA
<213> Human endogenous retrovirus

<400> 41
tgtgggtgga ggattaccca ggtgccgagg caagagactg aaggcacaaa ctgtttcagt 60
ataataaaga aaatggttag aataagaata gtcataatac aaattagata tagagatgat 120
catggacaat tatcaatcat tattataaac attattaatc attagctttt aatattactc 180
tttgttgcat tactaatata acctaggaat aaccggtggg tatagggtca ggtgctgaaa 240
ggacattggg agaagtgacc tagaaggcaa gaggtgagtc ttctgtcacg cccgcataag 300
ggttgcttga gggctccttg gtcaagtggt aacgccggtg tctgggaagg cacctgttac 360
ttagccgacc acgaaaggga gtctcctttc cttggaggag tcagggcgca ctctgctcca 420
ccagcttctt gtggaaggct ggatattatc caggcctgcc cgcagtcatc cggaggccta 480
aacccctccc tgtggtgctg tgcttcaatg ggcacactcc tcgtccactt tcatgttcct 540
cccatactcc tggtttctct ttgaagttcg tagtagatag tggtagaagg aatagggaaa 600
atcttaaagt gtttgatctt tcttataagt gcatagaaga aaacgctgac atatgctgcc 660
ttctctgtct gcttcagcta cctaagaggg aagggccccc tgtccagtga tcacgtgact 720
tgcttcacct tgtcaatcac ttagaagatt caccctcctt accctgcccc cttgtcttgt 780
atgcaataaa tatcagtgca cccagccttt cggggccact taccggtctc cacgtcttgg 840
tggtagtggt cccccgggcc cagctgtttt ctctttatct ctttgtcttg tgtcttattt 900
attacaatct ctcgtctccg cacacaggga gaacacccgc taagctccgt agggctggac 960
cctaca 966

<210> 42
<211> 398
<212> DNA
<213> Human endogenous retrovirus

<400> 42
aaacccctcc ctgtggtgct gtgcttcaat gggcacactc ctcgtccact ttcatgttcc 60
tcccatactc ctggtttctc tttgaagttc gtagtagata gtggtagaag gaatagggaa 120
aatcttaaag tgtttgatct ttcttataag tgcatagaag aaaacgctga catatgctgc 180
cttctctgtc tgcttcagct acctaagagg gaagggcccc ctgtccagtg atcacgtgac 240
ttgcttcacc ttgtcaatca cttagaagat tcaccctcct taccctgccc ccttgtcttg 300
tatgcaataa atatcagtgc acccagcctt tcggggccac ttaccggtct ccacgtcttg 360
gtggtagtgg tcccccgggc ccagctgttt tctcttta 398

<210> 43
<211> 938
<212> DNA
<213> Human endogenous retrovirus

<400> 43
tgtgggtgga ggattaccca ggtgccgagg caagagactg aaggcacaaa ctgtttcagt 60
ataataaaga aaatggttag aataagaata gtcataatac aaattagata tagagatgat 120
catggacaat tatcaatcat tattataaac attattaatc attagctttt aatattactc 180
tttgttgcat tactaatata acctaggaat aaccggtggg tatagggtca ggtgctgaag 240
ggacattggg agaagtgacc tagaaggcaa gaggtgagtc ttctgtcacg cccgcataag 300
ggttgcttga gggctccttg gtcaagtggt aacgccggtg tctgggaagg cacctgttac 360
ttagccgacc acgaaaggga gtctcctttc cttggaggag tcagggcaca ctctgctcca 420
ccagcttctt gtggaaggct ggatattatc caggcctgcc cgcagtcatc cggaggccta 480
aacccctccc tgtggtgctg tgcttcaatg ggcacactcc tcgtccactt tcatgttcct 540
cccatactcc tggttcctct ttgaagttcg tagtagatag tggtagaagg aatagggaaa 600
atcttaaagt gtttgatctt tcttataagt gcatagaaga aaacgctgac atatgctgcc 660
ttctctgtct gcttcagcta cctaagaggg aagggccccc tgtccagtga tcacgtgact 720
tgcttcacct tgtcaatcac ttagaagatt caccctcctt accctgcccc cttgtcttgt 780
atgcaataaa tatcagtgca cccagccttt cggkkcactt accggtctcc acgtcttggt 840
ggtagtggtc ccccggccca gctgttttct ctttatctct ttgtcttgtg tcttatttat 900
tacaatctct cgtctccgca cacagggaga acacccgc 938

<210> 44
<211> 396
<212> DNA
<213> Human endogenous retrovirus

<400> 44
aaacccctcc ctgtggtgct gtgcttcaat gggcacactc ctcgtccact ttcatgttcc 60
tcccatactc ctggttcctc tttgaagttc gtagtagata gtggtagaag gaatagggaa 120
aatcttaaag tgtttgatct ttcttataag tgcatagaag aaaacgctga catatgctgc 180
cttctctgtc tgcttcagct acctaagagg gaagggcccc ctgtccagtg atcacgtgac 240
ttgcttcacc ttgtcaatca cttagaagat tcaccctcct taccctgccc ccttgtcttg 300
tatgcaataa atatcagtgc acccagcctt tcggkkcact taccggtctc cacgtcttgg 360
tggtagtggt cccccggccc agctgttttc tcttta 396

<210> 45
<211> 963
<212> DNA

<213> Human endogenous retrovirus

<400> 45
tgtgggcgaa agattaccta ggtgccgagg caagagactg aaggcacaaa ctgtttcagt 60
ataataaaga aaatagttaa aataagaata gttataatac aaattagata tagagatgat 120
catggacaat tatcaatcat tattataaac attaatcatt agcttttaat attactcttt 180
gttgctttac taatataacc taggaataac cggtgggtat agggtcaggt gttgacggga 240
tattgtgaga agtgacctag aaggcaagag gtgagccttc tgtcacgccc acataagggc 300
cgcttgaggg ctctttggtc aagtggtaac gccagtgtct gtgaaggcac ctgttactta 360
gcagaccgcg aaagggagtc tcctttcctt ggaggagtca gggaacactc tgctccacca 420
gcttcttgtg gaaggctgga tattatctag gcctgcccgc agtcatctgg aggcctaaac 480
ccctccctgt ggtgctgtgc ttcagtggtc actctccttg tccactttca tgttcctccc 540
gtactcctgg ttcctctttg aagttcgtag tagatagcag tagaagaaat agtgaaagtc 600
ttaaagtatt tgatctttct tataagtgca tagaagaaaa cgctgacata tgctgccttc 660
tctatctctg cggtggctac ctaaaaggga agggcccccc gtcccatgat catgtgactt 720
gcttcacctt atcacttaga agattcatcc tccttaccct gcgccccctc gtcttgtatg 780
caataaatat cagcacgccc agtcgtttga ggccactgcc ggtctccgcg tcttggtggt 840
agtggtcccc cgggcccagc tattgtctct ttatctcttt gtcttgtgtc tttatttatt 900
acaatctctt gtctctgcac acagggagaa cacctgctaa gccccgtagg actggaccct 960
aca 963

<210> 46
<211> 397
<212> DNA
<213> Human endogenous retrovirus

<400> 46
aaacccctcc ctgtggtgct gtgcttcagt ggtcactctc cttgtccact ttcatgttcc 60
tcccgtactc ctggttcctc tttgaagttc gtagtagata gcagtagaag aaatagtgaa 120
agtcttaaag tatttgatct ttcttataag tgcatagaag aaaacgctga catatgctgc 180
cttctctatc tctgcggtgg ctacctaaaa gggaagggcc ccctgtccca tgatcatgtg 240
acttgcttca ccttatcact tagaagattc atcctcctta ccctgcgccc cctcgtcttg 300
tatgcaataa atatcagcac gcccagtcgt ttgaggccac tgccggtctc cgcgtcttgg 360
tggtagtggt ccccccgggcc cagctattgt ctcttta 397

<210> 47
<211> 489
<212> DNA
<213> Human endogenous retrovirus

<400> 47
tgttcaattc tttgccttct acttttaaac ttaacttcct cataaagcaa ccttttttcaa 60
tcacctgctc cactctgact cattctgatc acctgctcca ccctgactca ttccgatcac 120
ctgatccact gtgactcatt ccgattaccc gctccaccct gactcattct gattctgatt 180
tcctgctctg ccataaccat ttttcccgcc aaaccactca ccctgtcact ctctttaaat 240
tagccaattg gaattagttt agcctgtgcg gtctaaccct agccaatagg ggactgacac 300
agcagcaggg gccacatgtg tcaggaataa gacccccttc ccctccctgt ccagatgtgt 360
gctcaccatt gctccatctg tgagggcaca cccttctata gaagtaaatt gccttgctga 420
gaagaaaaaa aagaacattt tatattcaag tcctatttct tttgctgcac cgaaacttta 480
tttataaca 489

<213> Human endogenous retrovirus

<400> 45
tgtgggcgaa agattaccta ggtgccgagg caagagactg aaggcacaaa ctgtttcagt 60
ataataaaga aaatagttaa aataagaata gttataatac aaattagata tagagatgat 120
catggacaat tatcaatcat tattataaac attaatcatt agcttttaat attactcttt 180
gttgctttac taatataacc taggaataac cggtgggtat agggtcaggt gttgacggga 240
tattgtgaga agtgacctag aaggcaagag gtgagccttc tgtcacgccc acataagggc 300
cgcttgaggg ctctttggtc aagtggtaac gccagtgtct gtgaaggcac ctgttactta 360
gcagaccgcg aaagggagtc tcctttcctt ggaggagtca gggaacactc tgctccacca 420
gcttcttgtg gaaggctgga tattatctag gcctgcccgc agtcatctgg aggcctaaac 480
ccctccctgt ggtgctgtgc ttcagtggtc actctccttg tccactttca tgttcctccc 540
gtactcctgg ttcctctttg aagttcgtag tagatagcag tagaagaaat agtgaaagtc 600
ttaaagtatt tgatctttct tataagtgca tagaagaaaa cgctgacata tgctgccttc 660
tctatctctg cggtggctac ctaaaaggga agggcccccc gtcccatgat catgtgactt 720
gcttcacctt atcacttaga agattcatcc tccttaccct gcgccccctc gtcttgtatg 780
caataaatat cagcacgccc agtcgtttga ggccactgcc ggtctccgcg tcttggtggt 840
agtggtcccc cgggcccagc tattgtctct ttatctcttt gtcttgtgtc tttatttatt 900
acaatctctt gtctctgcac acagggagaa cacctgctaa gccccgtagg actggaccct 960
aca 963

<210> 46
<211> 397
<212> DNA
<213> Human endogenous retrovirus

<400> 46
aaacccctcc ctgtggtgct gtgcttcagt ggtcactctc cttgtccact ttcatgttcc 60
tcccgtactc ctggttcctc tttgaagttc gtagtagata gcagtagaag aaatagtgaa 120
agtcttaaag tatttgatct ttcttataag tgcatagaag aaaacgctga catatgctgc 180
cttctctatc tctgcggtgg ctacctaaaa gggaagggcc ccctgtccca tgatcatgtg 240
acttgcttca ccttatcact tagaagattc atcctcctta ccctgcgccc cctcgtcttg 300
tatgcaataa atatcagcac gcccagtcgt ttgaggccac tgccggtctc cgcgtcttgg 360
tggtagtggt cccccgggcc cagctattgt ctcttta 397

<210> 47
<211> 489
<212> DNA
<213> Human endogenous retrovirus

<400> 47
tgttcaattc tttgccttct acttttaaac ttaacttcct cataaagcaa ccttttttcaa 60
tcacctgctc cactctgact cattctgatc acctgctcca ccctgactca ttccgatcac 120
ctgatccact gtgactcatt ccgattaccc gctccaccct gactcattct gattctgatt 180
tcctgctctg ccataaccat ttttcccgcc aaaccactca ccctgtcact ctctttaaat 240
tagccaattg gaattagttt agcctgtgcg gtctaaccct agccaatagg ggactgacac 300
agcagcaggg gccacatgtg tcaggaataa gacccccttc ccctccctgt ccagatgtgt 360
gctcaccatt gctccatctg tgagggcaca cccttctata gaagtaaatt gccttgctga 420
gaagaaaaaa aagaacattt tatattcaag tcctatttct tttgctgcac cgaaacttta 480
tttataaca 489

29/pts

P12088

Retroviral expression vectors on the basis of HERV LTR sequences

The present invention relates to retroviral expression vectors bearing promoters which may be cell-specifically controlled. The vectors are useful for example for the cell-specific expression of genes of therapeutic value in the context of a gene therapy.

Retroviruses are RNA viruses wherein the viral genes are encoded by a single-stranded RNA molecule. Following entry of the viruses into the cell, the viral RNA is converted into a double-stranded DNA molecule by means of reverse transcription. The DNA enters the nucleus and integrates into the cellular chromosome. The integrated form of viral DNA, the so-called provirus, represents the template for the expression of viral genes.

Integration of the viral genome into the cellular chromosome is an obligatory step of viral replication and is mediated by virus-encoded enzymes. With few exceptions it appears that the viability of the infected cell is not or almost not affected by the presence of the retroviral genome in the cell, the expression of its genes and the formation of viral particles.

Retroviral gene transfer is used for the introduction of functional genes, in particular genes of therapeutic value, into cells without affecting the ability of the host cell to proliferate. Due to their mode of replication retroviruses are suitable for such gene transfer. In the most simple

embodiment, at least a portion of the viral genes is replaced by a gene of interest and, by using the efficient viral infection process, this gene of interest is transferred into the target cell.

Retroviral vectors are suitable for gene therapy because the infection by retroviruses occurs with high efficiency and the retroviral vectors may be modified to incorporate heterologous DNA and may stably integrate into the genome of the host cell. A plurality of retroviral vectors has been developed in recent years, and by way of example reference is made herein to the reviews by Günzburg et al. (1996) and Robbins et al. (1998).

A possible preferred embodiment for retroviral vectors are so-called ProCon vectors which have been described for the first time in WO 96/07748. For the disclosure reference is made to this document in its entirety.

ProCon vectors have heterologous promoter elements and optionally other regulatory elements in their 3' LTR which following infection are duplicated and translocated to the 5' LTR in the target cell and which are capable of controlling the expression of marker genes or therapeutic genes. These heterologous genes are not directly linked to the promoter but are inserted inside the vector.

ProCon vectors comprise an 5' LTR portion having the structure U3, R, U5, and at least one coding and/or non-coding sequence as well as a 3' LTR region comprising a U3 portion which is completely or partially deleted wherein the deleted U3 portion has been replaced by a polylinker sequence followed by the R and U5 portions.

Propagation of these vectors is performed by means of a helper cell line producing high amounts of viral proteins which are no longer synthesized by the expression vector itself. However, the helper cell line is no longer capable of producing a replication-competent virus. This cell line is also referred to as packaging cell line and comprises a cell line transfected with at least a second plasmid carrying the genes which enable the packaging of the modified retroviral vector. In this respect, reference is made to WO92/10564 which is incorporated herein by reference in its entirety.

The DNA encoding the modified retrovirus (expression vector) is transfected into the packaging cell line. Under these conditions the modified retroviral genome comprising the inserted therapeutic genes or marker genes, respectively, is transcribed and packaged into retroviral particles (recombinant viral particles). Then, this recombinant virus is used for the infection of target cells; the genome of the modified retrovirus, i.e. the expression vector, is integrated into the target cell genome, wherein this integration occurs together with the marker genes or therapeutic genes, respectively. A cell infected with the recombinant viral particles generated in this manner is unable to produce new vector virus since no other viral proteins are present in these cells. The DNA of the expression vector containing the genes of therapeutic value or marker genes, respectively, which has been integrated into the host cell is present in the cellular DNA in an integrated form and may subsequently be expressed in the cell.

Preferably, genes of therapeutic value for which an expression is achieved by means of such retroviral expression

vectors are to be expressed in a cell- and tissue-specific manner. For this purpose, cell-specific regulatory sequences are introduced into the LTR sequence of the expression vector. For example, these cell-specific regulatory sequences comprise cell-specifically controllable promoter regions, cell-specific enhancer sequences as well as binding sites for transcription factors. The promoters are localized in the U3 portion of the LTR.

Presently, cellular promoter sequences or promoter sequences of exogenous retroviruses are inserted into retroviral vectors.

Cellular promoters often require additional signal structures which may be present at a great distance upstream or downstream of the promoter. Therefore, it has always been found difficult to isolate strong, tissue-specific, cellular promoter sequences and to clone them into retroviral vectors. Promoters of exogenous retroviruses bear the advantage that they contain, within the retroviral LTR, all the necessary regulatory elements in a confined region and, therefore, may be transcribed essentially independent of neighboring DNA sequences at the site of integration. However, a severe disadvantage is that although they are strong, they are not tissue-specific and generally are expressed with equal strength in all cell types.

Therefore, it is an object of the present invention to provide novel retroviral expression vectors which utilize the benefits of retroviral promoters to concentrate all signal structures required for transcription in a confined region within the U3 and R regions, but simultaneously avoid the disadvantages associated therewith.

According to the present invention, this object has been achieved by inserting into retroviral expression vectors a cell-specifically controllable promoter portion derived from a human endogenous retroviral DNA nucleotide sequence (HERV). These promoter sequences of human endogenous retroviral viruses are already present in the host cell, and it has been found according to the invention that they are excellently suitable for regulation of the cell-specific expression of marker genes and genes of therapeutic value.

Endogenous retroviruses (ERV) may be found in the genome of all cells in an organism. They are transferred vertically via the germ line and may be reactivated by conditions caused by the environment. About 2% of the human genome consist of endogenous retroviruses and retroviral sequences, solitary HERV LTRs being present in an amount of 20,000-40,000 copies per genome (Tab. 1) (Leib-Mösch et al., 1993; Wilkinson et al., 1994; Patience et al., 1997).

Since HERV sequences have been integrated into the primate genome already 30-40 millions of years ago, it may be assumed that in the course of evolution most of the pathogenic sequences were eliminated from the provirus by mutations and rearrangements or have been modified, respectively, to be no longer disadvantageous for the organism. Compared to vectors derived from animal viruses, retroviral vectors constructed from such sequences have the advantage that no new viral sequences must be introduced into the genome. In addition, also by recombination with HERV sequences already present in the genome no novel retroviruses may arise as it might be the case if retroviruses of other species were used as vectors. For this reason, the use of these sequences in the

construction of retroviral vectors can minimize the safety risk. Furthermore, homologous regions contained in the genome may be utilized for a tissue-specific integration of the retroviral vectors into specific sites of a chromosome.

In the course of evolution, HERV elements have adopted a number of cellular functions. For example, promoter and enhancer elements of HERV LTRs are used for the transcriptional control of cellular genes (Kato et al., Feuchter-Murthy et al., 1993; Di Christofano et al., 1995). One example for the use of LTR regulatory elements for a tissue-specific expression of a cellular gene is the human amylase gene. This gene is controlled by the LTR of an HERV-E element and in this manner is restricted specifically to be only expressed in saliva glands (Ting et al., 1992). Moreover, Schulte and co-workers (1996) have shown that the insertion of an endogenous retrovirus into the 5' untranslated region of the pleiotrophin gene is responsible for the throphoblast-specific activity thereof (Schulte et al., 1996). In other instances, polyA signals of HERV LTRs may also serve to polyadenylate cellular transcripts (Mager, 1989; Goodchild et al., 1992).

Since retroviruses must maintain their transcriptional activity most independently of the surrounding regions of their site of integration, the main advantage of the use of retroviral promoters resides in the fact that all signal structures required for transcription are localized in a confined region within the U3 region and the R region of the LTR, as described above. Because these HERV promoters have persisted in the primate genome since millions of years they have adapted during evolution to be active in a cell type-

specific manner similar to cellular promoters and thus combine the advantages of cellular and retroviral promoters.

HERVs are transcribed starting from a classical RNA polymerase II promoter (Wilkinson et al., 1994). This promoter is localized within the LTR region. Therefore, the HERV transcript comprises no complete copy of the provirus. To compensate for the loss of transcription control elements, these elements have developed the mechanism of reverse transcription by which the lost sequences at both ends of the elements are regenerated from which in turn the LTRs are regenerated. Besides promoter sequences the HERV LTRs also contain a plurality of different binding sites for transcription factors (Seifarth et al., 1998) responsible for the tissue-specificity of expression.

Although there is a certain structural similarity between HERVs and exogenous animal retroviruses, such as MLV or MMTV, HERV sequences and in particular HERV promoters have never been considered as possible candidates for the development of retroviral expression vectors. In contrast, up to now they have only been considered as disruptive factors in view of gene therapy (Patience et al., 1997). Because of sequence homologies, there has been concern that they might interfere with the therapeutic vector by recombination in the target cell. Although it was impossible up to now to confirm these concerns by experimentation, however, a problem arose in the development of very efficient human packaging cell lines that co-packaging and inadvertent transfer of potentially infectious HERV sequences occurred. Therefore, a detailed study of the possible packaging of expressed HERV sequences into virions based on MLV has been conducted. Patience et al. (1998) identified mRNA transcripts of several different HERV

families, such as HERV-K and HERV-H, in human packaging cell lines. Even by using a highly sensitive RT-PCR test, however, none of these sequences could be detected in the MLV vector particles released by the cells.

According to these findings, a packaging and transfer of HERV sequences and therefore, eventually, also of HERV-based vectors in MLV packaging systems seemed to be out of question. Although HERV genes have a sequence homology of 50-65% with respect to MLV genes, particularly the regions which are essential for packaging and infection, and in particular the packaging signal localized between the 5' LTR and the gag region as well as the LTR itself have no detectable sequence homology with respect to the corresponding MLV sequences. Since up to now there are no cell lines known which produce HERV particles in sufficient amounts, however, presently, efficient HERV packaging systems are also not conceivable.

Thus, the present invention solves the problem of retroviral expression vectors controlling the cell- and tissue-specific expression of foreign genes (gene of interest) by providing expression vectors containing, in functional assembly, at least the following elements:

a) DNA sequences for packaging of the vector RNA and for the cell-specific expression of proteins or peptides encoded by heterologous DNA nucleotide sequences;
b) one or more heterologous DNA nucleotide sequences (transcription unit) encoding a protein or peptide;
wherein the DNA sequences for cell-specific expression are characterized by comprising a cell-specifically controllable promoter region derived from a human endogenous retroviral virus, in particular from the LTR sequence of said virus.

The promoter region of the HERV sequence may comprise the whole LTR region of the HERV. However, in another embodiment of the present invention, the promoter region only comprises the U3 region or the R-U3 region of the HERV LTR. In another preferred embodiment of the present invention, besides these regions the promoter region also comprises the untranslated region between the 5' LTR and the gag genes. It has been found according to the present invention that this region also contains sequences which control the cell-specific expression of proteins or peptides, respectively, i.e. which at least contribute to said cell-specific expression.

The promoters are partial regions of DNA required for the start of transcription of the corresponding structural genes. The promoter includes the transcription start site, the recognition and binding site for RNA polymerase. The promoter may also comprise other sequences to which regulatory proteins may bind and which thereby specifically control the initiation of transcription. Examples of such proteins are transcription factors and repressors. Examples of said regulatory elements of the transcription activity are the CAAT box, GC box and TATA box. The promoters are recognized by a type II polymerase.

The promoter regions for the cell-specific expression of foreign proteins from HERVs may optionally be combined with other sequences derived from exogenous retroviruses which promote cell-specific expression. In addition, there is considered a combination with regulatory sequences from cellular genes to support cell-specific expression.

Furthermore, the retroviral expression vector according to the present invention at least contains DNA sequences for packaging of the vector by means of a packaging helper cell line. The DNA sequences for packaging are localized between the 5' LTR and the gag gene. Such packaging signals are present in any retroviral vector and therefore known to the skilled artisan. Examples of packaging signals are listed in Mann et al., 1985, and Rein, 1994, as well as the literature cited therein. This literature is incorporated by reference in its entirety.

The retroviral expression vector according to the present invention contains one or more transcriptional units encoding an amino acid sequence. The amino acid sequence refers to a protein or peptide. Any sequence encoding a protein or peptide of interest may be inserted into the expression vector. For example, such proteins or peptides may be encoded by marker genes, genes of therapeutic value, genes with antiviral function, anti-tumor genes and/or cytokin genes. This list could be continued to any number. The genes which can be introduced into the retroviral expression vector are known to those skilled in the art. The type of genes inserted depends on the intended use of the vector according to the present invention.

For example, the vectors according to the present invention may be employed for gene therapy to transfer heterologous DNA into target cells in order to render diseases accessible to a specific therapy. The vector DNA is introduced into the selected target cell so that the heterologous DNA is expressed in the target cell and the product encoded by the DNA is produced. This includes particularly such genes for the expression of proteins which are not produced or not

produced any longer or not produced in sufficient amounts by the target cell so that a disease condition develops. The invention not only comprises such proteins or peptides, respectively, which occur naturally but also those which have been modified in a manner to achieve a desired effect, for example a higher enzyme activity, blocking of a binding site for viruses, destruction of tumor cells by suicide genes, etc.

Generally, the DNA nucleotide sequences encoding a protein or a peptide is heterologous DNA encoding RNA and proteins which the cell in which the proteins or peptides, respectively, are expressed usually does not produce in vivo. It may also be referred to as foreign DNA. This includes any protein, such as enzymes, hormones, and antibodies. Therefore, the retroviral expression vectors provided by the present invention are designed to express proteins of interest in human cells.

The promoter regions employed according to the present invention are selected from HERV sequences derived from the HERV families known. Examples of these are HERV-K, HERV-H, HERV-E, HERV-L, HERV-T, HERV-R, HERV-I, HERV-P, ERV9, HERV-W.

It should be understood that it is also possible to screen other, presently unknown HERV families in order to find promoter sequences which are presently unknown and which regulate the cell-specific expression.

Preferred LTR sequences from HERVs according to the present invention which may be employed for the tissue-specific expression of proteins and peptides are disclosed in the annex. They may be introduced into retroviral expression

vectors to achieve the object according to the present invention. It should be understood, however, that by means of methods known per se it is also possible to select only portions of these LTRs to keep the sequences inserted into the vector as small as possible. Useful fragments may be selected using various deletion mutants. Furthermore, also other variations of these LTR sequences are possible, e.g. point mutations, insertions, additions, substitution of several nucleotides, etc. in order to increase the efficiency of the tissue-specific expression and to adapt it to the desired function.

In a preferred embodiment according to the present invention the ProCon vectors described at the beginning are employed. Such ProCon vector comprise a 5' LTR region having the structure U3-R-U5, one or more sequences encoding a protein or peptide and optionally non-coding sequences as well as a 3' LTR portion comprising a partially or completely deleted U3 region wherein the deleted U3 portion at least comprises the HERV LTR sequences employed according to the present invention, followed by an R-U5 region. Further details are described for example in WO96/07748 and WO96/28564. These documents are included herein by reference in their entirety.

According to the present invention a strategy has been developed to track down promoter sequences having a cell-specific function. This strategy is described in more detail in the following description. It has to be understood that principally also other methods for the finding of HERV LTR sequences which act in a cell-specific manner may be considered and used. Thus, the present invention is not limited to the following examples.

The retroviral expression vectors according to the present invention are packaging deficient, i.e. are unable to produce viral particles without assistance by a packaging helper cell line. Therefore, the present invention also comprises a retroviral vector system containing a retroviral expression vector as described in the present invention and a packaging cell line which contains at least a retroviral or recombinant retroviral construct encoding the packaging proteins of the retroviral expression vector. Such packaging cell lines are known per se and have been described. By way of example, reference is made herein to the murine packaging cell line PA317 (Saller et al., 1998).

In the following, the invention will be described in general, followed by a description with respect to Examples.

According to the present invention, the applicability of human endogenous retroviruses for the development of tissue-specific vectors for gene therapy has been investigated. For this purpose, first the tissue-specificity of HERV pol transcription has been examined in different cell lines, such as T cells, keratinocytes and breast cancer cells using a "reverse dot blot" procedure. In this test, the expression pattern of the various HERV families was found to be generally cell type dependent. To isolate HERV LTRs with transcriptional activity from different cell lines and tissues, primers were developed which could be used for the specific amplification of the U3/R regions from mRNA preparations. The isolated LTR sequences as well as individual members of already known LTRs were inserted into expression vectors. Following transient transfection of the reporter plasmids, the activity of the LTR promoters was tested in the different cell lines via the luciferase

activity or eGFP fluorescence, respectively. The promoter activity of individual HERV LTRs was found to vary clearly dependent on the cell line tested. The promoter region of a HERV-H LTR isolated from astrocytes and liver cells which was found to be especially active in lung fibroblast cells (LC5) in several tests was inserted into two retroviral promoter conversion vectors (pLESN and pLX), tested in packaging cell lines, the packaging efficiency was evaluated, and after infection of the target cell was tested for the occurrence of a promoter conversion. FACS analyses were performed to detect the transcriptional activity in the target cells.

Thus, a method has been described by which HERV promoter sequences (U3/R region) mediating a tissue-specific expression may be identified and isolated. Subsequently, the tissue-specificity and promoter activity of these sequences was tested in a transient transfection assay in various human cell lines. Eventually, suitable sequences were chosen, cloned into a promoter conversion vector (ProCon vector) whereby their usefulness for the construction of tissue-specific vectors for gene therapy was examined. The preparation of the retroviral expression vectors according to the present invention is carried out using recombinant techniques known per se. Such techniques are for example described in Sambrook et al., 1989, and Perbal, 1984. For the construction of the ProCon vectors see WO 96/07748 already mentioned at the beginning an the related literature.

3. Results

3.1 Analysis of HERV transcription in different cell types

To investigate HERV transcription in different cells a method (reverse dot blot hybridization) was employed in the first step which had been originally developed for the detection of HERV expression in peripheral blood mononucleated cells (Herrmann and Kalden, 1994). For this method, cloned and characterized HERV *pol* gene fragments from human genomic DNA were immobilized on a membrane and hybridized with radiolabeled HERV *pol* gene probes. The probes were amplified from mRNA of different cells using RT PCR and degenerated oligonucleotides homologous to a highly conserved region of retroviral *pol* genes (Shih et al., 1989; Donehower et al., 1990). Using this method we obtained a characteristic hybridization pattern with every cell line examined so far which was the first result to indicate a tissue-specific expression of HERV elements.

3.2 Isolation of LTR U3 regions of expressed HERVs

The tissue-specific expression of a retrovirus is primarily defined by its U3 region. In this region, all regulatory sequences are localized, such as promoter, enhancer, and binding sites for various cellular transcription factors. For this reason, primers were developed which could be used for the specific isolation by means of RT PCR of these HERV sequences from the mRNA of different cell lines (Tab. 2; Fig. 1). In this manner, about 30 different HERV LTRs were cloned. In part, these sequences were tested in a reporter plasmid for their promoter activity and tissue-specificity.

In a first approach, for the PCR a polydT primer was combined with a primer complementary to the polypurine stretch (PPT) of retroviral RNA (Fig. 1). The PPT stretch is a conserved portion in the non-translated region between the *env* gene and

the U3 region of the 3' LTR. During reverse transcription of the retrovirus, the PPT region is used as a primer binding site for plus strand synthesis (Sorge and Hughes, 1982).

By means of data base analyses, the PPT sequences of different HERV families were identified and classified into different groups by comparing their homology. From the consensus sequences of individual groups oligonucleotides were synthesized as primers for RT PCR. The mRNA was prepared from different cell lines: epithelial cells (HeLa, HaCaT), fibroblast cells (LC5), T cells (H9, HUT78), lymphoblasts (CML), glioma cells (85HG66, U373), pancreatic cells (MiaPaCa2, Panc1), liver cells (Chang Liver), and breast cancer cell lines (T47-D, MCF7). Moreover, cDNA libraries (Clontech) of various human tissues (brain, heart, liver, kidney, lung, pancreas, placenta, skeletal muscle) were also employed in the RT PCR.

Subsequently, the fragments obtained were cloned, sequenced, and analyzed by means of data base comparison. Among the PCR fragments obtained with PPT and polydT primers two LTRs were assigned to the families of HERV-H and HERV-K due to a comparison of homologies. By using polydT primers in these PCR samples, numerous sequences were amplified which did not reveal any homologies to known retroviral LTRs and moreover did not contain any promoter structural elements. For this reason, other sequences were selected for primer synthesis from conserved regions of the U3 region and from the R region of the HERV-K and HERV-H families (Mold et al., 1997) (Fig. 1, Tab. 1). The resulting PCR products were separated on an agarose gel, followed by transfer to nitrocellulose filters and hybridization with probes prepared from the LTR regions of different HERV LTRs (HERV-K-pl167, HERV-H-H6, HERV-E,

HERV-L). Afterwards, the hybridizing fragments were cloned into a vector (pZERO, Invitrogen) and sequenced. Using this method, several HERV LTRs could be isolated which are listed in Table 3.

The HERV-K LTRs isolated from human brain and heart tissue as well as from T47-D cells show very strong sequence homologies to the 3' LTR of HERV-K10. In contrast, the HERV-H LTRs exhibited much higher sequence variations. HERV-H31, HERV-H3, HERV-HCM1, HERV-HCM4, HERV-HMP23 are homologous to the HERV-H-H6 LTR isolated by Mager et al., the other HERV-H sequences show homologies to the HERV-H LTRs from vervet monkey, marmoset and man isolated by Anderssen et al. (1997). The HERV-W LTRs isolated from T47-D cells are related to the LTR of clone CL6 (Komurian-Pradel, 1989).

3.3 Analysis of the expression of HERV promoters in a transient luciferase assay

For an analysis of the promoter activity and tissue-specificity of the isolated HERV LTRs, these were first cloned into a luciferase reporter plasmid (pBL, Butz, K., DKFZ, Heidelberg). This vector contains the luciferase gene of *Photinus pyralis* fused to the SV40 polyA signal of pBLCAT2 (Hoppe-Seyler et al., 1991).

The individual vector constructs were transiently transfected into different cell lines. After 48 h, the luciferase activity from the cell lysate was measured using the luciferase assay kit of Promega company and was determined as the relative luciferase activity after standardization for ß-galactosidase activity or *Renilla* luciferase activity, respectively. The LTR promoter activities were determined in

epithelial cells (HeLa, HaCaT), fibroblast cells (LC5), T cells (H9, HUT78), glioma cells (85HG66, U373), liver cells (Chang Liver), pancreatic cells (MiaPaCa2, Panc1), and breast cancer cell lines (T47-D, MCF7).

The results are presented in Figures 2a-2f. According to these results, among all endogenous LTRs tested the HERV-H-H6 LTR has the strongest promoter. The HERV-K LTR from placenta is particularly active in HeLa cells. In all other cell lines, this LTR exhibits only a very weak activity. Also in HeLa cells, HERV-K-T47-D showed a strong activity, this LTR was also active in HaCaT cells and pancreatic cells. The HERV-L LTR has a strong promoter activity in liver cells and a weak activity in T cells and pancreatic cells. The HERV-T-S71A and HERV-E LTRs were active in none of the cell lines tested. Also, no activity at all of a HERV LTR could be observed up to now in CML cells.

Almost all of the cloned HERV-H LTRs (HERV-H1, HERV-H8, HERV-H13, HERV-H19, HERV-H H6, Tab. 3) were active in 85HG66 cells while HERV-H1 and HERV-H8 showing the highest activity in this cell line (not shown). HERV-H19 was very active in HeLa cells. The HERV-HCM1 LTR exhibited the highest promoter activity in all cell lines and was especially active in lung fibroblasts (LC5) (Fig. 3).

3.4 Construction of HERV hybrid vectors and monitoring of the activities of HERV promoters in these vectors

The functionality of human endogenous retroviral LTR sequences in retroviral vectors was tested in two different promoter conversion vectors (ProCon). For this purpose, hybrid HERV/MLV vectors were constructed using two vectors

pLESN-MMTV (Fig. 7) and pLX-MMTV (Fig. 8) on the basis of MLV. These vectors include the EGFP gene as a reporter gene which is expressed from the 5' LTR (in varying amounts depending whether measured before or after the promoter conversion) as well as a neomycin gene which is expressed from an SV40 promoter. Moreover, vector pLX-MMTV contains a prokaryotic origin of replication enabling recloning of the provirus for further molecular characterizations.

To construct the HERV hybrid vectors, in each case the MMTV LTR was replaced by the HERV-HCM1 LTR (Fig. 7). For this purpose, the LTR was amplified first by means of PCR from the vector pBL-HERV-H using specific primers which contained additional sequences for the restriction enzymes MluI and SacII. Then, these fragments were inserted into the vectors having their 3' U3 deleted. After transfection into the packaging cell line, the EGFP reporter gene is first expressed from the MLV promoter (Fig. 9a). After infection of the target cells and successful promoter conversion by reverse transcription in the target cells, the reporter gene is present under the transcriptional control of the HERV LTR.

The HERV hybrid vector constructs pLESN-HERV-H (Fig. 7) and pLX-HERV-H (Fig. 8) as well as the parent vectors pLESN-MMTV and pLX-MMTV were transfected into the amphotrophic packaging cell line PA317. Afterwards, the resulting retroviral vector particles were used for the infection of cell lines CrfK and LC5.

The infected cell lines were cloned and the selected cellular clones were examined for the presence of vector constructs and for the occurrence of promoter conversion. For this purpose, chromosomal DNA was prepared from infected and

uninfected cells and analyzed by means of PCR. The primers were selected from the MLV U3 (P5) and R (P2) region as well as the HERV-H region (P1) and used for the PCR in combination with a primer for the EGFP region (Fig. 9a). The PCR products were hybridized with HERV-H-specific probes (Fig. 9b). After amplification with the primers P1 and P3, the DNA infected by pLX HERV-H particles yielded a PCR product of 1.1 kb which hybridized to the HERV-H probe. Amplification using MLV U3-specific primers (P2/P3) with DNA of cells infected with pLX and pLX HERV-H gave PCR products having a size of about 900 bp which showed not hybridization to the HERV-H probe. No PCR product hybridizing to the HERV-H probe was obtained from the amplification using MLV R primers (P5/P3). These results show that promoter conversion has occurred and that the MLV promoter of the 5' LTR had been replaced by the HERV promoter.

After integration into the target cell DNA, the HERV LTR promoter activity in the retroviral vectors was determined by FACS analyses via the measurement of the EGFP fluorescence (Fig. 10). For this purpose, the activity of the starting vector pLX-MMTV was compared with that of the HERV vector pLX-HERV-H (H6) prior to and after induction with dexamethasone. The vector containing the MMTV LTR may be activated by dexamethasone. The vector containing the HERV LTR is not activated by dexamethasone, however, its activity is by a factor of 10 higher as compared to the dexamethasone-stimulated MMTV hybrid vector.

3.5 Effect of regulatory elements in the R and U5 regions on the promoter activity of HERV sequences

In order to examine which sequence region is required for a functional HERV promoter, the effect of additional LTR sequences localized outside of the U3 region in the LTR was investigated in several examples. For this purpose, the activity in the luciferase assay of the U3 region of 7 HERV-K LTRs (HERV-K-T47D, L5, L50, L8, L9, L48, and L20/49) was compared to the activity of the corresponding U3-R fragments. It was surprisingly found that the different R regions are able to affect the promoter in the U3 region in a very different manner. In the group 1 LTRs (L5, L50, L8, L9) the presence of R sequences resulted in a marked increase in promoter activity in all cell lines tested (Fig. 4a). In contrast, in the group 2 LTRs (L20/L49) the HERV promoter activity is reduced by the R region (Fig. 4b). The HERV-K-T47D promoter (Fig. 5) and the L48 promoter (not shown) are substantially unaffected by the respective R sequences. Interestingly, in the case of the HERV-K-T47D LTR sequence regions localized downstream of the U3-R region and comprising the U5 region as well as the 3' non-translated region and the start of the *gag* gene have a clearly activating effect (Fig. 5).

A sequence analysis of the different R regions tested revealed that group 1 LTRs have a binding site for transcription factor SP1 in the R region which is missing from the R region in group 2 LTRs (Fig. 6). In contrast, the group 2 R region contains a potential binding site for factor TFS3 which acts as a repressor of transcription. This shows that the activity of HERV promoters may be modified by insertion of additional regulatory elements such as transcription factor binding sites, enhancer sequences, or negative-regulatory elements.

References

Anderssen, S., Sjottem, E., Svineng, G., Johansen, T. (1997). Comparative analyses of LTRs of the ERV-H family of primate-specific retrovirus-like elements isolated from marmoset, african green monkey, and man. Virology **234**. 14-30

Di Christofano, A., Strazzullo, M., Longo, L., La Mantia, G. (1995). Characterization and genomic mapping of the ZNF80 locus: expression of this zinc-finger gene is driven by a solitary LTR of ERV9 endogenous retroviral family. Nucl. Acid. Res. **23**. 2823-2830

Donehower L. A., Bohannon, R. C., Ford, R. A. (1990). The use of primers from highly conserved *pol* regions to identify uncharacterized retroviruses by the polymerase chain reaction. J. Virol. Methods **28**. 33-46

Emmerman, M., Temin, M. H. (1986). Comparision of promotor suppression in avian and murine retrovirus vectors. Nucl. Acid. Res. **14**. 9381-9396

Feuchter, A., Mager, D. (1990). Functional heterogeneity of a large family of human LTR-like promoters and enhancers. Nucl. Acid. Res. **18**. 1261-1270

Feuchter-Murthy, A. E., Freeman, J. D., Mager, D. L. (1993). Splicing of a human endogenous retrovirus to a novel phospholipase A2 related gene. Nucl. Acid. Res. **21**. 135-143

Goodchild N. L., Wilkinson, D. A., Mager, D. L. (1992). A human endogenous long terminal repeat provides a polyadenylation signal to a novel, alternatively spliced transcript in normal placenta. Gene, **121**. 287-294

Herrmann, M., Kalden, J. R. (1994). PCR and reverse dot hybridisation for the detection of endogenous retroviral transcripts. J. Virol. Methods **46**. 333-348

Hoppe-Seyler, F., Butz, K., zur Hausen, H. (1991). Repression of the Human Papillomavirus Type 18 Enhancer by the Cellular Transcription Factor Oct-1. J. Virol. **65**. 5613-5618

Kato-N., Shimotohno, K., VanLeeuwen, D., Cohen, M. (1990). Human proviral mRNAs downregulated in choriocarcinoma encodes zinc finger protein related to Kruppel. Mol. Cell. Biol. **10**. 4401-4405

Komurian-Pradel, F., G. Paranhos-Baccala, F. Bedin, A. Ounanian-Paraz, M. Sodoyer, C. Ott, A. Rajoharison, E. Garcia, F. Mallet, B. Mandrand, and H. Perron, 1999. Molecular cloning and characterization of MSRV-related sequences associated with retrovirus-like particles. Virology 260: 1-9.

Leib-Mösch, C., Haltmeier, M., Werner, T., Geigl, E.-M., Brack-Werner, R., Francke, U., Erfle, V., Hehlmann, R. (1993). Genomic distribution and transcription of solitary HERV-K LTRs. Genomics **18**. 261-269

Mager, D. L. (1989). Polyadenylation Function and Sequence Variability of the Long Terminal Repeats of Human Endogenous Retrovirus-like Family RTVL-H. Virology **173**. 591-599

Mann, R., and D. Baltimore. 1985, Varying the position of a retrovirus packaging sequence results in the encapsidation of both unspliced and spliced RNAs. J. Virol. 54:401-407.

Mold, D. E., Wu, T., Askin, F., Huang, R. C. (1997). Four classes of HERV-K long terminal repeats and their relative promoter strengths for transcription. Biomed. Science **4**. 78-82

Perbal, B., 1984. A Practical Guide to Molecular Cloning. John Wiley & Sons.

Price, J., Turner, D., Cepko, C. (1987). Lineage analysis in the vertebrate nervous system by retrovirus-mediated gene transfer. Proc. Natl. Acad. Sci. **89**. 9237-9241

Rein, A., 1994, Retroviral RNA packaging: a review. Arch. Virol. Suppl. 9:513-22.

Saller, R. M., Öztürk, F., Salmons, B., Günzburg, W. (1998). Construction and characterization of a hybrid mouse mammary tumor virus/murine leukemia virus-based retroviral vector. J. Virol. **2**. 1699-1703

Salmons, B., Günzburg, W. H. (1993). Targeting of retroviral vectors for gene therapy. Human Gene Therapy **4**. 129-141

Sambrook, J., E. F. Fritsch, and T. Manatis. 1989. Molecular Cloning. Cold Spring Harbor Laboratory Press, New York, USA.

Schulte, A. M., Lai, S., Kurtz, A., Czubayko, F., Riegel, A. T. (1996). Human trophoblast and choriocarcinoma expression of the growth factor pleiotrophin attributable to germ-line insertion of an endogenous retrovirus. Proc. Natl. Acad. Sci. **93**. 14759-14764

Seifarth, W., Baust, C., Murr, A., Skladny, H., Krieg-Schneider, F., Blusch, J., Werner, T., Hehlmann, R., Leib-Mösch, C., (1998). Proviral Structure, chromosomal location and expression of HERV-K-T47D, a novel human endogenous retrovirus derived from T47-D particles. J. Virol. **72**. *In press*

Shih, A., Misra, R., Rush, M. G. (1989). Detection of multiple, novel reverse transcriptase coding sequences in human nucleic acids: relation to primate retroviruses. J. Virol. **63**. 64-75

Sorge, J., Hughes, S. H. (1982). Polypurine tract adjacent to the U_3 region of the Rous sarcoma virus genome provides a *cis*-acting function. J. Virol. **43**. 482-488

Ting, C.-N., Rosenberg, M. P., Snow, C. M., Samuelson, L. C., Meisler, M. H. (1992). Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. Genes Dev. **6**. 1457-1465

Wilkinson, D. A., Mager, D. L., Leong, J.-A. (1994). Endogenous human retroviruses. In "The retroviridae" Vol. **3** (J. Levy, ed.), pp. 465-535, Plenum Press, NY.

Patience, C., Wilkinson, D. A., Weiss, R. A. (1997). Our retroviral heritage. Trends Genet. 13, 116-120.

Patience, C., Takeuchi, Y., Cosset, F.-L., Weiss, R. A. (1998). Packaging of endogenous retroviral sequences in retroviral vectors produced by murine and human packaging cells. J. Virol. 72, 2671-2676.

26

Amended Claims

1. Retroviral expression vector containing at least the following elements in functional assembly:
a) DNA sequences for packaging of the vector RNA and for cell-specific expression of proteins or peptides encoded by heterologous DNA nucleotide sequences;
b) one or more DNA nucleotide sequences encoding a protein or peptide
wherein said DNA sequences for the cell-specific expression contain a cell-specifically controllable promoter region from a human endogenous retroviral DNA nucleotide sequence (HERV).

2. Expression vector according to claim 1 wherein said DNA sequences for cell-specific expression are derived from the LTR region and optionally from the non-translated region between the 5' LTR and the gag region of HERVs.

3. Vector according to claim 1 or claim 2 wherein the whole LTR region, the U3 region, or the R and U3 regions are derived from a human endogenous retroviral nucleotide sequence.

4. Vector according to one or more of the preceding claims wherein said nucleotide sequences encoding one or more proteins or peptides are selected from one or more elements of the group consisting of marker genes, therapeutic genes, antiviral genes, anti-tumor genes, and cytokin genes.

5. Vector according to one or more of the preceding claims wherein said cell-specifically controllable promoter region is derived from the LTR region of a cell-specifically expressed endogenous human retroviral nucleotide sequence.

6. Vector according to one or more of the preceding claims wherein said human endogenous retroviral cell-specifically controllable promoter sequences are selected from one or more promoter sequences of HERV families of the group consisting of HERV-K, HERV-H, HERV-E, HERV-L, HERV-T, HERV-R, HERV-I, HERV-P, ERV9, HERV-W.

7. Vector according to one or more of the preceding claims wherein said promoter region besides the TATA box additionally comprises recognition and binding sites for regulatory proteins.

8. Vector according to claim 7 wherein said recognition and binding sites for regulatory proteins comprise the GC box, the CAAT box, enhancer sequences and repressor sequences as well as hormone responsive sequence motifs and wherein, optionally, additional recognition and binding sites for regulatory proteins from the LTR region of exogenous retroviruses and/or from cellular genes are comprised.

9. Vector according to one or more of the preceding claims wherein said vector is a promoter conversion vector comprising a 5' LTR portion having the structure U3-R-

U5, one or more sequences selected from coding and non-coding sequences, and a 3' LTR portion comprising a U3 region which is partially or completely deleted wherein the deleted U3 portion is replaced by a cell-specifically controllable promoter region from a HERV LTR sequence, followed by the R-U5 region.

10. The mRNA or RNA of a retroviral expression vector according to one or more of the preceding claims.

11. Prokaryotic cell or eukaryotic cell containing a retroviral expression vector according to one or more of the preceding claims.

12. Eukaryotic cell containing a retroviral expression vector according to one or more of the preceding claims in an integrated form.

13. Use of a cell-specifically controllable promoter region from a human endogenous retroviral DNA nucleotide sequence for the regulation of the expression of foreign genes in retroviral expression vectors, preferably in ProCon vectors.

14. Use of an expression vector according to one or more of the preceding claims for the expression of foreign genes in gene therapy.

15. Virion containing a retroviral expression vector RNA obtained by transcription of an expression vector DNA according to one or more of the preceding claims.

16. Method for the preparation of a virion according to claim 14 for the introduction of one or more nucleotide sequences encoding a protein or peptide wherein said retroviral expression vector according to one or more of the preceding claims is introduced into a suitable packaging cell line under such conditions that the virion is formed and released by the packaging cell line.

17. Method for the introduction of nucleotide sequences encoding one or more proteins or peptides into an eukaryotic cell wherein said cell is infected by a virion as defined in claim 14 under such conditions that the nucleotide sequences encoding the protein or peptide is inserted into the chromosomal DNA of the eukaryotic cell.

18. Method according to claim 17, wherein the eukaryotic cell is a mammalian cell.

19. Process according to claim 18, wherein the mammalian cell is a human cell.

20. Retroviral vector system comprising a retroviral expression vector according to one or more of the preceding claims and a packaging cell line comprising at least one retroviral or recombinant retroviral construct encoding for the packaging proteins of the retroviral expression vector.

21. Retroviral vector system according to claim 20, wherein the packaging cell line comprises retroviral or recombinant retroviral constructs encoding for such retroviral proteins which are not encoded by the retroviral expression vector.

SUMMARY

The present invention relates to retroviral expression vectors with cell-specifically controllable promoters. For example, the vectors may be used for the cell-specific expression of genes of therapeutic value in the context of a gene therapy.

The present invention describes retroviral expression vectors containing at least the following elements in functional assembly:

a) DNA sequences for the packaging of the vector RNA and for cell-specific expression of proteins or peptides encoded by heterologous DNA nucleotide sequences;

b) one or more DNA nucleotide sequences encoding a protein or peptide wherein said DNA sequences for the cell-specific expression thereof contain a cell-specifically controllable promoter region from a human endogenous retroviral DNA nucleotide sequence (HERV).

P12088

Applicant: GSF-Forschungszentrum für Umwelt und Gesundheit GmbH

Title: Retroviral expression vectors on the basis of HERV LTR sequences

Description of the Figures of the PCT application

Tab. 1: Human endogenous retroviral elements with indication of the family, copy number and percentage in the genome

Tab. 2: Primers used for the amplification of different HERV LTR regions

Tab. 3: HERV LTRs analyzed
A: HERV LTRs from different cell lines and tissues indicating the homology
B: HERV LTRs published in the literature

Fig. 1: RT PCR strategy for the isolation of the U3/R regions of transcribed HERV sequences

Fig. 2g: Relative promoter activities of different HERV LTRs in various cell lines

Fig. 3a: Cell line plotted vs. relative promoter activity

Fig. 3b: Cell line plotted vs. relative promoter activity

Fig. 4: LTR R region modulated/promoter activity of HERV-K-T47D-related LTRs
A: Activation, cell line plotted vs. relative promoter activity
B: Inhibition, cell line plotted vs. relative promoter activity

Fig. 5: Sequences downstream of LTR R modulate the promoter activity of HERV-K-T47D-related LTRs
Cell line plotted vs. relative promoter activity

Fig. 6: Regulatory elements in the R region of HERV-K T47D LTRs

Fig. 7: Retroviral ProCon vectors pLESN-MMTV and pLESN-HERV-H

Fig. 8: Retroviral ProCon vectors pLX-MMTV and pLX-HERV-H

Fig. 9: a) Promoter conversion of ProCon hybrid vectors
b) Detection of correct promoter conversion by means of PCR and hybridization using a HERV-H and a psi probe

Fig. 10: a) Organization of the two ProCon vectors pLX-MMTV and pLX-HERV-H
b) Promoter activity of the HERV-H LTR as compared to the MMTV LTR following infection of CrfK cells

Tab. 1: Human endogenous retroviral elements

	HERV family	Copy number	% of genome
Class I HERVs (type C- related HERVs)	**HERV-ERI**		0.07 %
	HERV-E (4-1, ERVA, NP-2)	35 - 50	
	HERV-E LTR	500 - 600	
	51-1	35 - 50	
	ERV1	10 - 15	
	HERV-R (ERV3)	10	
	RRHERV-I	20	
	HERV-T (S71, CRTK1, CRTK6)	50 - 60	
	HERV-T LTR	150 - 200	
	ERV-FRD	8	
	HERV-RW		0.2%
	HERV-W (MSRV)	25 - 50	
	HERV-R (ERV9)	30 - 40	
	ERV9 LTR	3000 - 4000	
	HERV-P (HuERS-P, HuRRS-P)	50 - 90	0.01%
	HERV-IP		0.01%
	HERV-I (RTVL-I)	25 - 50	
	HERV-IP-T47D (ERV-FTD)	35	
	HERV-IP LTR	1800 - 2000	
	HERV-HF		0.2%
	HERV-H (RTVL-H, RGH)	900 - 1000	
	HERV-F	16	
	HERV-H-LTR	1000	
Class II HERVs (type A-, B-, and D-related HERVs)	**HERV-K**		0.5%
	HERV-K(HML-1)	10 - 20	
	HERV-K(HML-2)	30 - 50	
	HERV-K10		
	HERV-K-HTDV		
	HERV-K-IDDM		
	HERV-K(HML3)	25	
	HERV-K(HML-4)	6	
	HERV-K-T47D		
	HERV-K(HML-6)	30 - 40	
	HERV-K-HML-6p		
	HERV-K-HML-6.17		
	HERV-K(HML-7)	?	
	HERV-K-NMW7		
	HERV-K(HML-8)	?	
	HERV-K(HML-9)	?	
	HERV-K-NMW9		
	HERV-K(HML-10)	10 - 50	
	HERV-KC4		
	HERV-K LTR	10 000 - 25 000	
Foamy virus-related HERVs	**HERV-L**	100 - 200	0.02%

Tab. 2: Primers used for the amplification of different HERV-LTR-regions

Nr.	Primer	Sequence
34	HERV-K	ATGGCGGTTTTGTCGAA
35	HERV-K	GTTCCMTYAGTATTTATTGATC
36	HERV-K	ATGGAGCATACAATCGGG
3	HERV-K	AAGAAAAGGGGGAAATGTGGG
11	HERVKC4	AAAGGGAGGGGGGCATG
12	HERV-KT47-D	TAAAAAGGGGGGAGATG
1	HERV-H	ATGTGAGCAACATGGCTGTTTATTTC
2	HERV-H	TGTCAGGCCTCTGAGCCCAA
39	HERV-H	GCCATCTCGAGTGTCAGSCCTCTGAGYCYARGC
37	HERV-H	TATCTTGAATTCGKGTGAGCAAYAARRCTTTA
31	polydT	TTTTTTTTTTTTTTTT
17	HERV-E	AAAGGGGGGGAAATATG
18	HERV-L	AGGGGTGGGACTTGCGATG
19	HERV-W	TGTTGAGATGGGGGACTGAG
20	HERV-W	GCAGTTGCAAGATTTAATAGAG

Tab. 3: Analyzed HERV-LTRs

A: HERV-LTRs from different cell lines and tissues

	Primer	Herkunft	Homology
HERV-K2	34/36	T47-D	HERV-K10, M12854, 97,9 %
HERV-K3	34/36	T47-D	HERV-K10, M12854, 98,4 %
HERV-K22-K32-K27-K45	34/36	brain	HERV-K10, M12854, 98,6 % *
HERV-K30	3/31	heart	HERV-K10, M12854, 97,6 %
HERV-K-T47D-L5		T47-D	MRSV, AF127229
HERV-K-T47D-L50		T47-D	MRSV, AF127229
HERV-K-T47D-L8		T47-D	MRSV, AF127229
HERV-K-T47D-L9		T47-D	MRSV, AF127229
HERV-K-T47D-L48		T47-D	MRSV, AF127229
HERV-K-T47D-L20		T47-D	MRSV, AF127229
HERV-IP-T47D		T47-D	MRSV, AF127229
HERV-T47D-W2	19/20	T47-D	MRSV, AF127229
HERV-T47D-W4	19/20	T47-D	MRSV, AF127229
HERV-T47D-W5	19/20	T47-D	MRSV, AF127229
HERV-W1	19/20	T47-D	MRSV, AF127229
HERV-W10	19/20	T47-D	MRSV, AF127229
HERV-W11	19/20	T47-D	MRSV, AF127229
HERV-W18	19/20	T47-D	MRSV, AF127229
HERV-W2	19/20	T47-D	MRSV, AF127229
HERV-W22	19/20	T47-D	MRSV, AF127229
HERV-W23	19/20	T47-D	MRSV, AF127229
HERV-W4	19/20	T47-D	MRSV, AF127229
HERV-W5	19/20	T47-D	MRSV, AF127229
HERV-W6	19/20	T47-D	MRSV, AF127229
HERV-W8	19/20	T47-D	MRSV, AF127229
HERV-H1	1/2	H9	Cercopithecus aethiops ERV-H; U96012, 87,1%
HERV-H8	1/2	HUT	HERV-H LTR18106, 84,8%
HERV-H13	1/2	HUT	HERV-H LTR18106, 91,8%
HERV-H19	1/2	liver	Callithrix jacchus ERV-H, 5´LTR; U96052, 92,1%
HERV-H31	1/2	liver	HERV-H(H6) x12717, 99,8%
HERV-H3	1/31	85HG66	HERV-H(H6) x12717, 100 %
HERV-H CL1	1/2	Chang Liver	HERV-H(H6) x12717, 100 %
HERV-H CL2	1/2	Chang Liver	HERV-H LTR18106, 84 %
HERV-H CL3	1/2	Chang Liver	Callithrix jacchus ERV-H, 5´LTR Silva 5, U96057, 84,2 %
HERV-H CL4	1/2	Chang Liver	HERV-H(H6) x12717, 100 %
HERV-H PA7	1/2	Panc1	Callithrix jacchus ERV-H, 5´LTR Silva 4, U96062, 85,7 %
HERV-H PA8	1/2	Panc1	Cercopithecus aethiops ERV-H, Vero 22, U96012, 87,1%
HERV-H PA9	1/2	Panc1	HERV-H LTR18106, 85 %
HERV-H PA10	1/2	Panc1	Callithrix jacchus ERV-H, 5´LTR Silva 4, U96062, 85,6 %
HERV-H MC14	1/2	MCF7	Cercopithecus aethiops ERV-H, Vero 22, U96012, 86,6%
HERV-H MC15	1/2	MCF7	Cercopithecus aethiops ERV-H, U96012, 86,6 %
HERV-H MC16	1/2	MCF7	Callithrix jacchus ERV-H, 5´LTR Silva 4, U96062, 87,4 %
HERV-H MC17	1/2	MCF7	Cercopithecus aethiops ERV-H, Vero 22, U96012, 86,6%
HERV-H MP20	1/2	MiaPaca	Human beta globin retrovirus-like repetitive element, k01891, 92,8 %
HERV-H MP21	1/2	MiaPaca	HERV-H LTR18106, 89,2 %
HERV-H MP23	1/2	MiaPaca	HERV-H(H6) x12717, 99,5 %

B: HERV-LTRs published in the literature

	(bp)	reference
HERV-K-pI167	970	Leib-Mösch *et al*, 1993
HERV-K-T47-D	1200	Seifarth *et al.*, 1998
HERV-H-H6	393	Feuchter und Mager, 1990
HERV-T-S71A	625	Murr, Dissertation, 1998
HERV-E	391	Steele *et al* , 1984
HERV-L	462	Cordonnier *et al.*, 1995

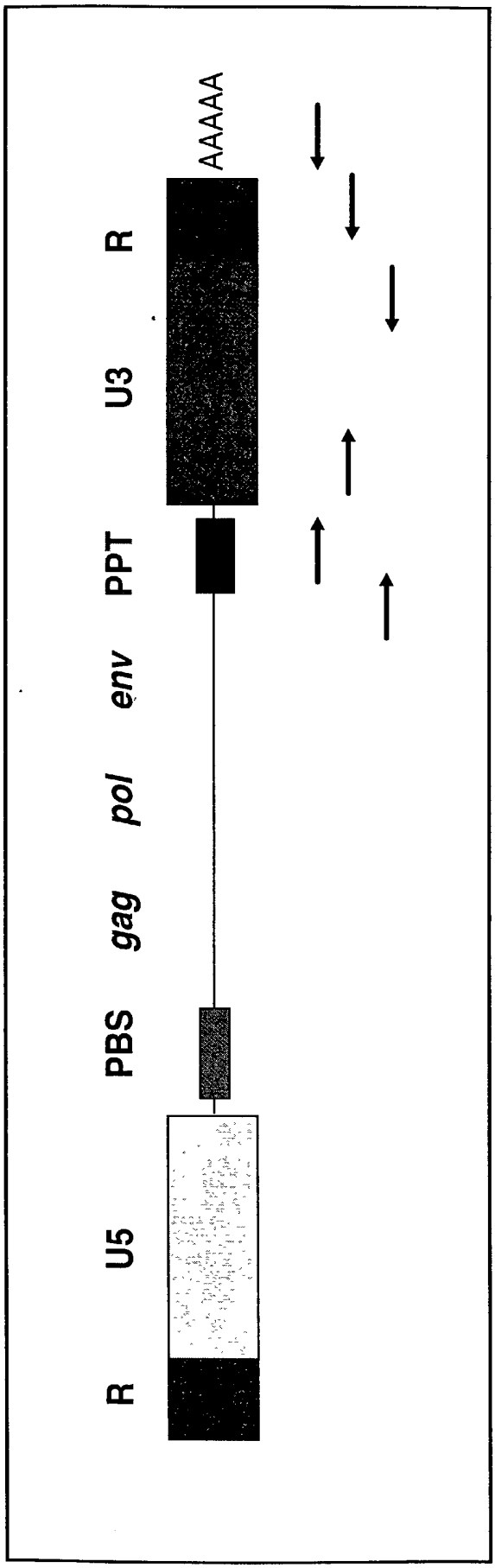


Fig.1: RT-PCR strategy to isolate U3/R-regions of transcribed HERVs


HERV-H-H6
90
80
70
60
50
40
30
20
10
0
MCF7
Hela
LC5
U373
MiaPa
HaCat
Panc1
Chang Liver
H9
HUT

Abb.2a)

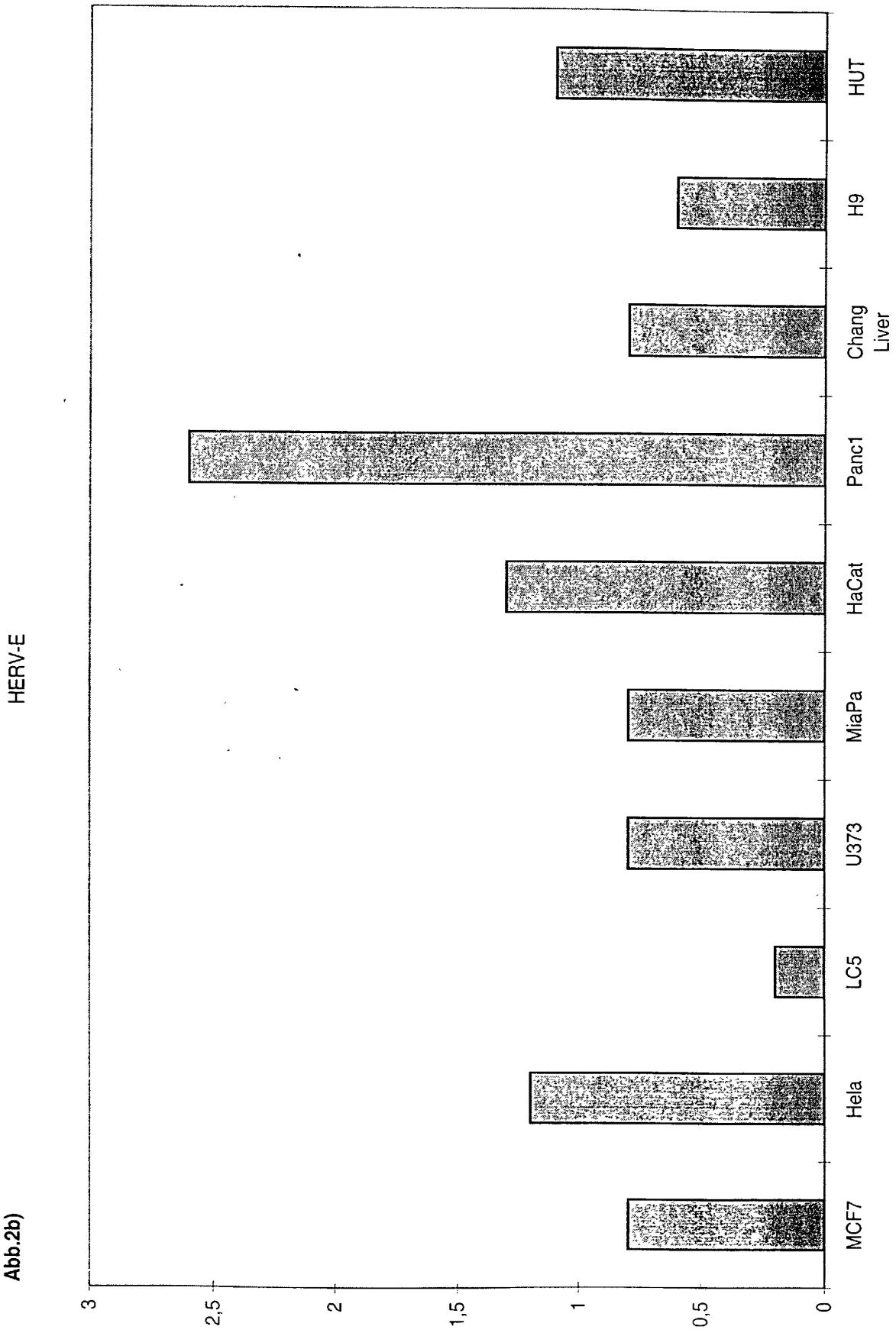

HERV-E
3
2,5
2
1,5
1
0,5
0
MCF7
Hela
LC5
U373
MiaPa
HaCat
Panc1
Chang Liver
H9
HUT

Abb.2b)

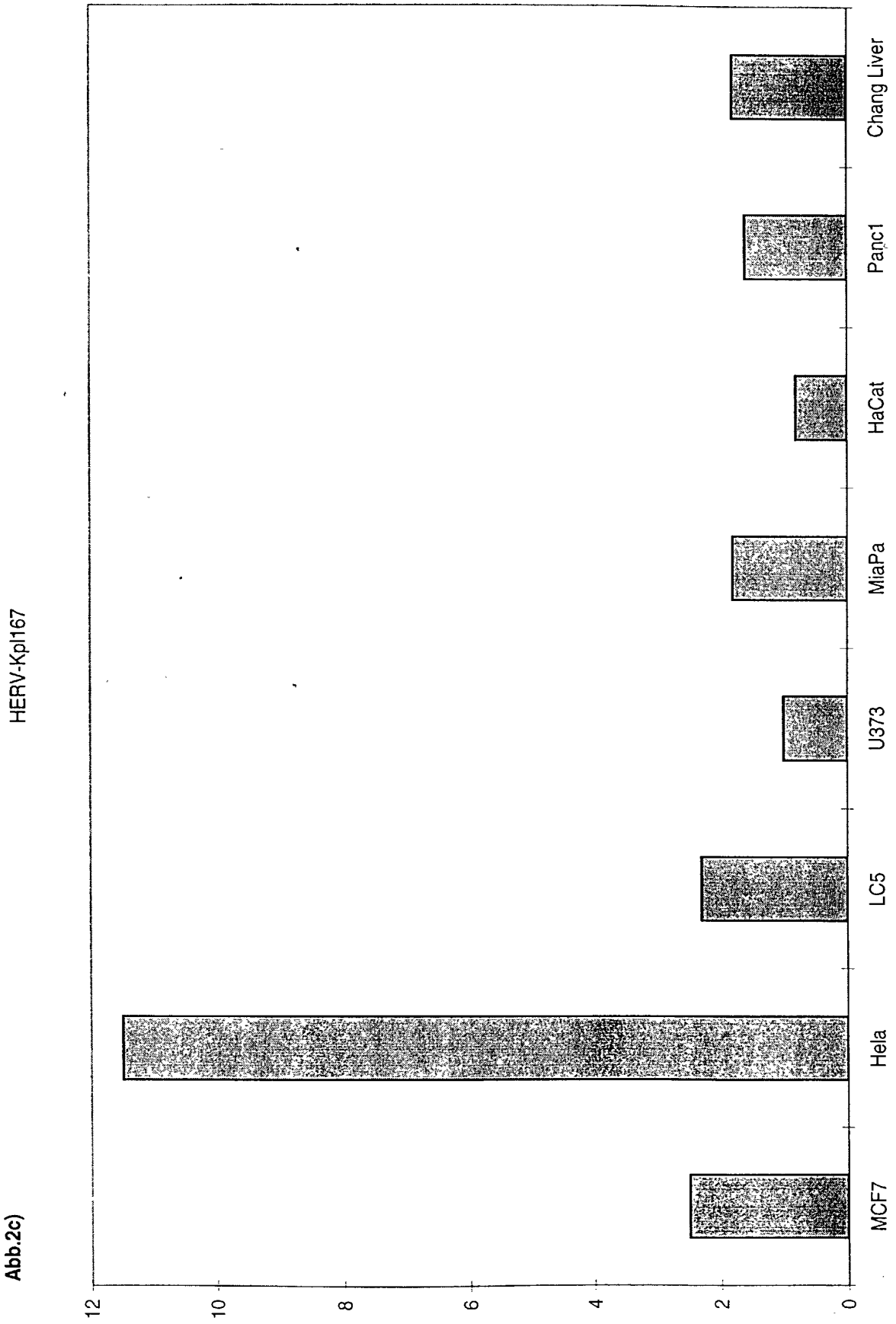

HERV-Kpl167
12
10
8
6
4
2
0
MCF7
Hela
LC5
U373
MiaPa
HaCat
Panc1
Chang Liver

Abb.2c)

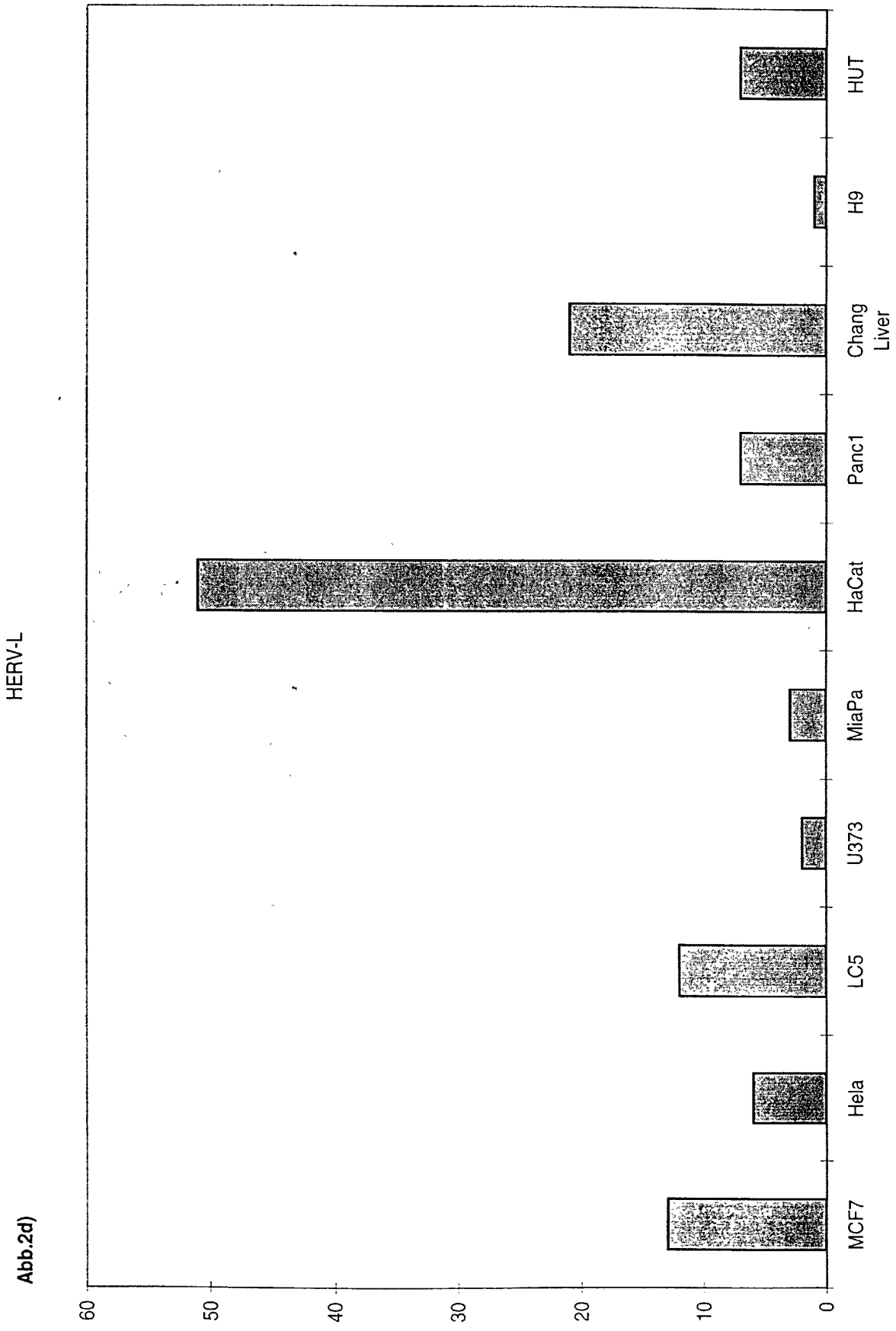

HERV-L
60
50
40
30
20
10
0
MCF7
Hela
LC5
U373
MiaPa
HaCat
Panc1
Chang Liver
H9
HUT

Abb.2d)

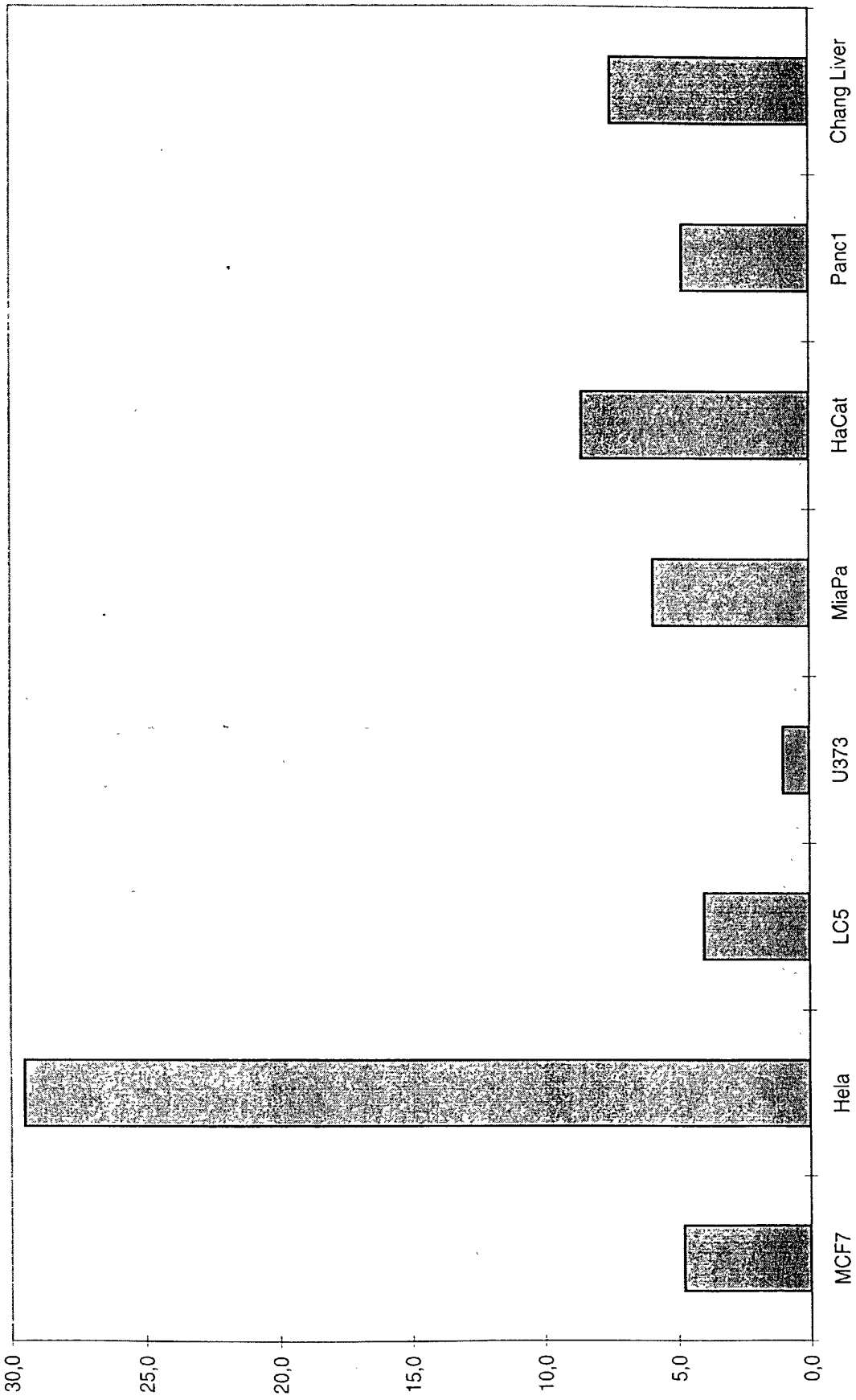

HERV-K-T47D
30,0
25,0
20,0
15,0
10,0
5,0
0,0
MCF7
Hela
LC5
U373
MiaPa
HaCat
Panc1
Chang Liver

Abb.2e)

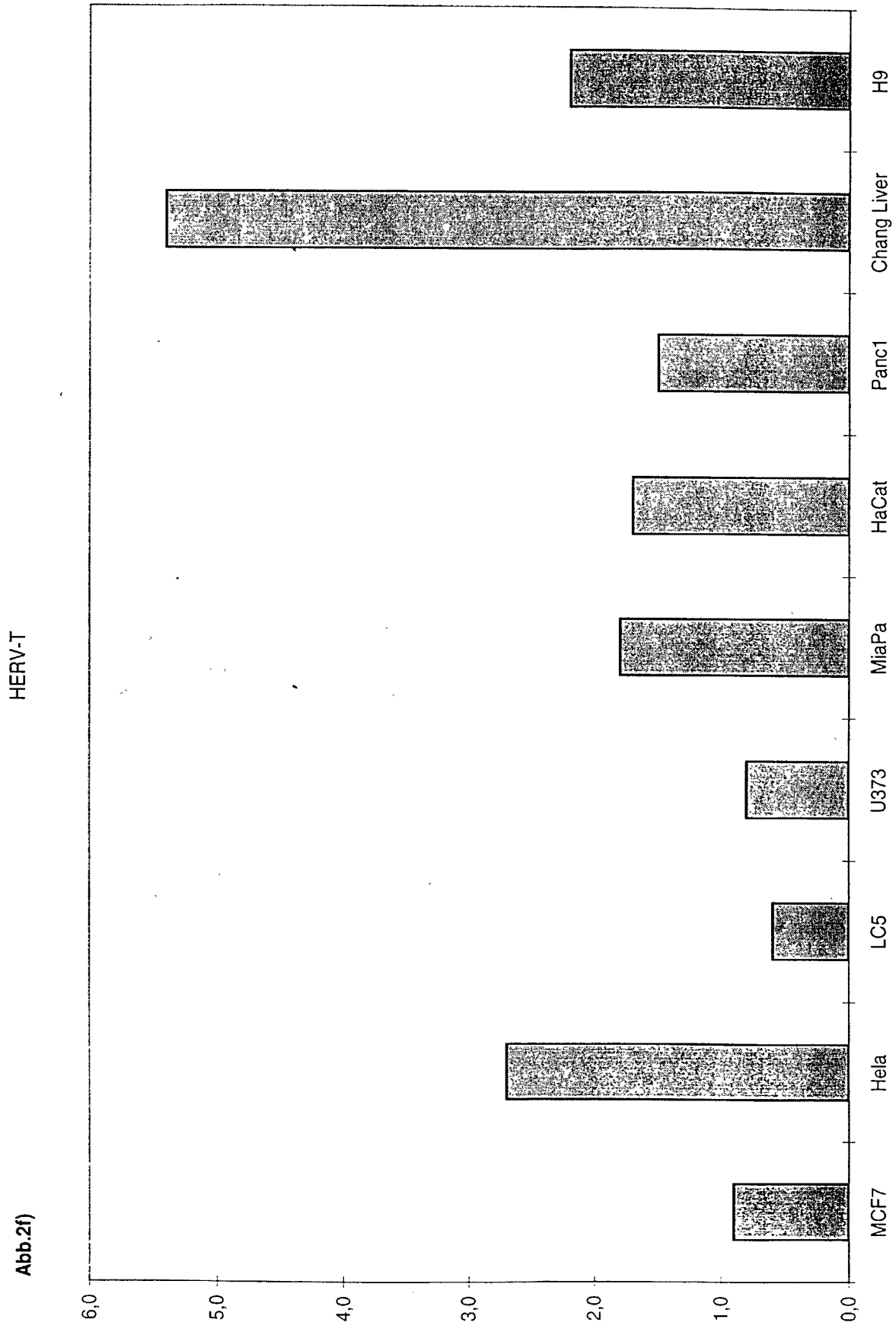

HERV-T
6,0
5,0
4,0
3,0
2,0
1,0
0,0
MCF7
Hela
LC5
U373
MiaPa
HaCat
Panc1
Chang Liver
H9

Abb.2f)

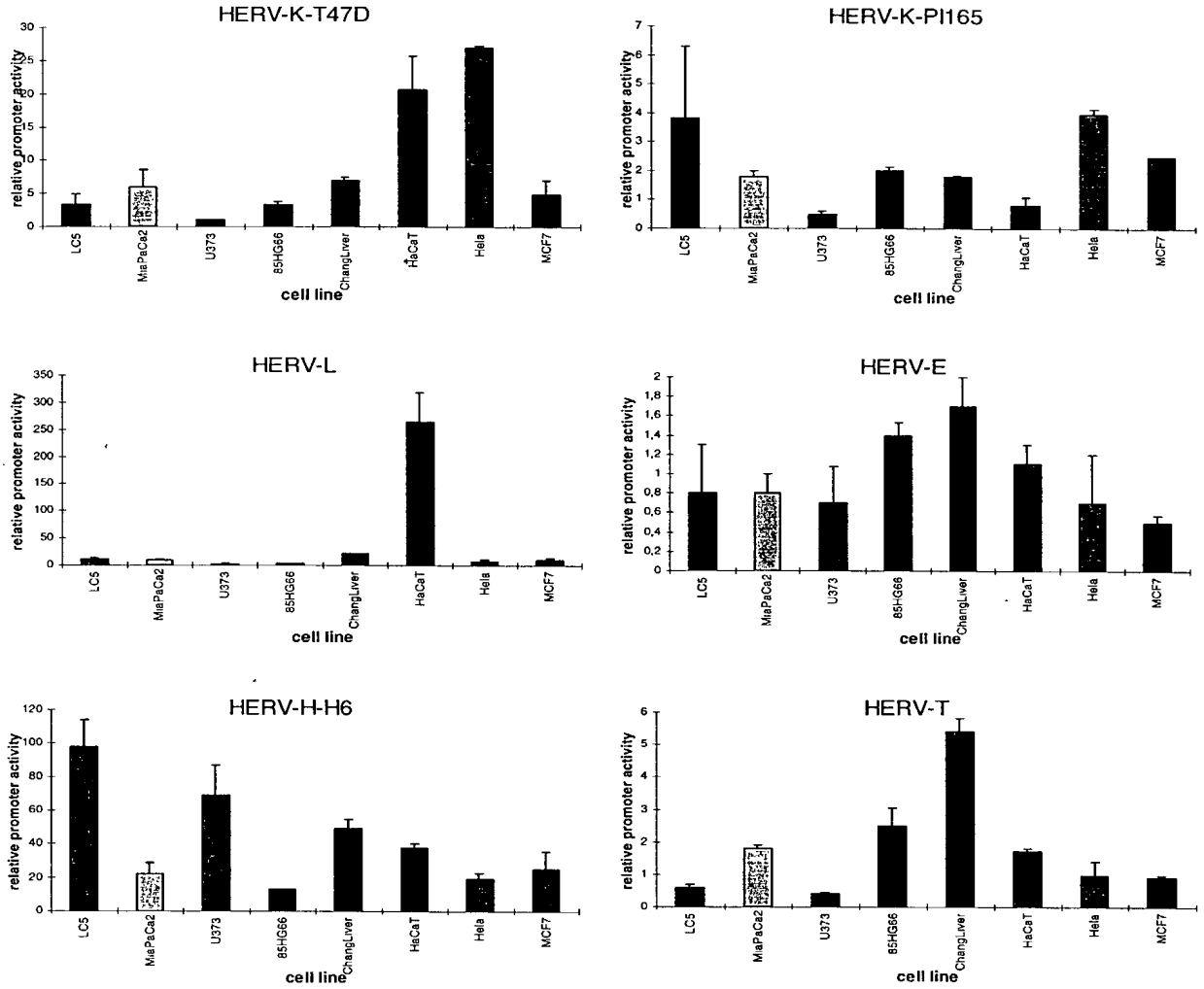
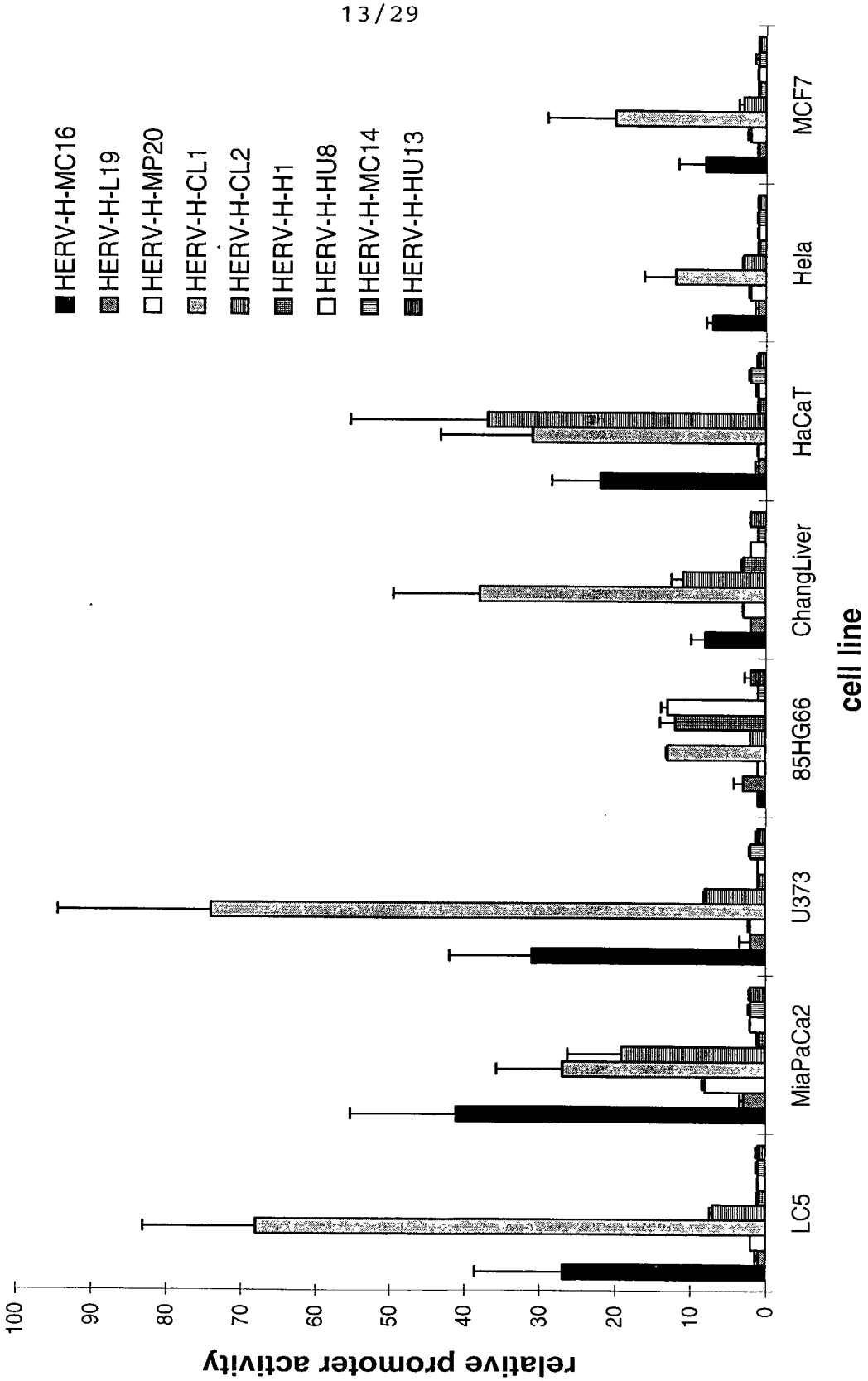


Abb. 2g: relative promoter activity of different HERV-LTRs in different cell lines

Fig. 3a


HERV-H-MC16
HERV-H-L19
HERV-H-MP20
HERV-H-CL1
HERV-H-CL2
HERV-H-H1
HERV-H-HU8
HERV-H-MC14
HERV-H-HU13
relative promoter activity
100
90
80
70
60
50
40
30
20
10
0
LC5
MiaPaCa2
U373
85HG66
ChangLiver
HaCaT
Hela
MCF7
cell line

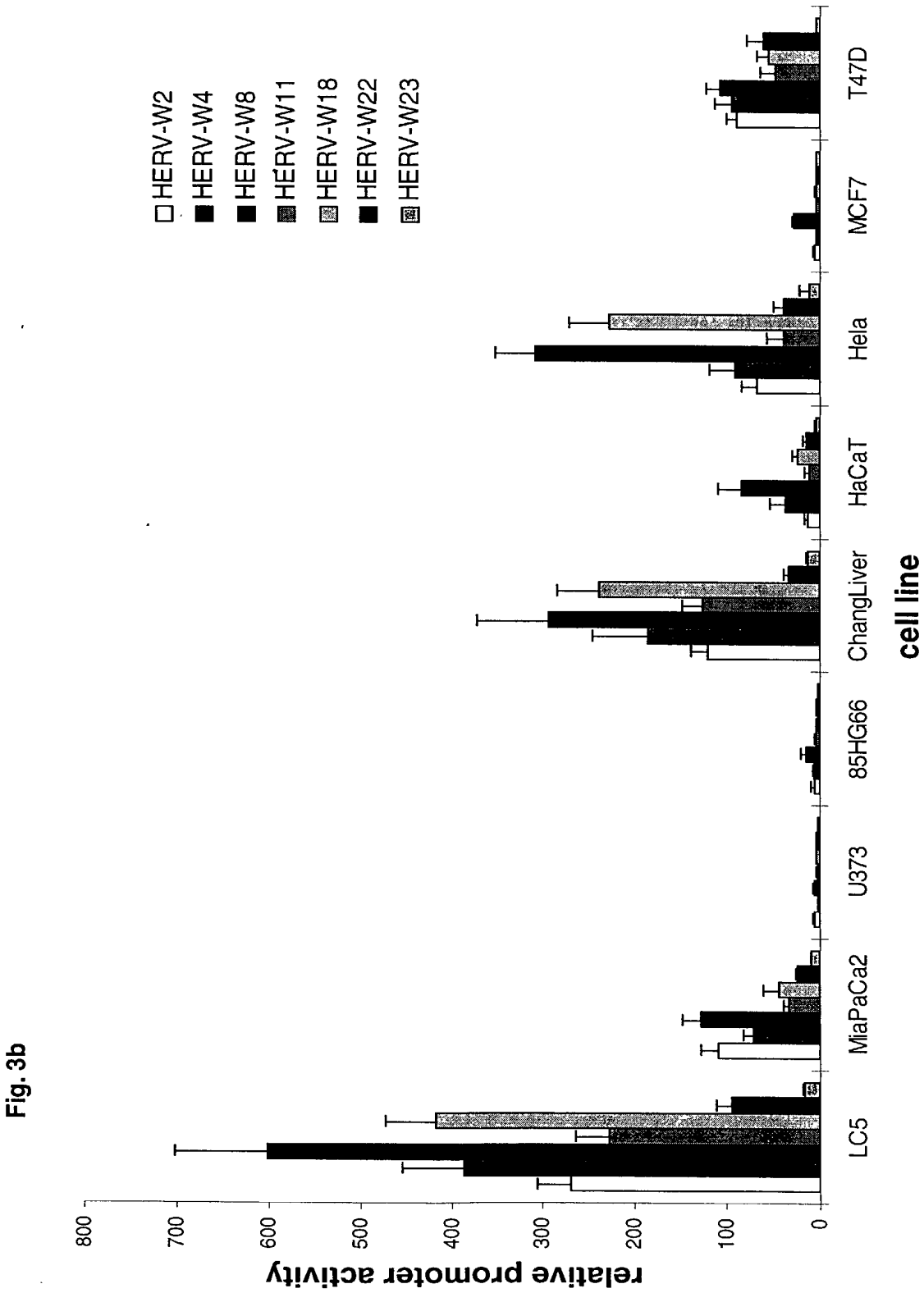

HERV-W2
HERV-W4
HERV-W8
HERV-W11
HERV-W18
HERV-W22
HERV-W23
relative promoter activity
0
100
200
300
400
500
600
700
800
LC5
MiaPaCa2
U373
85HG66
ChangLiver
HaCaT
Hela
MCF7
T47D
cell line

Fig. 3b

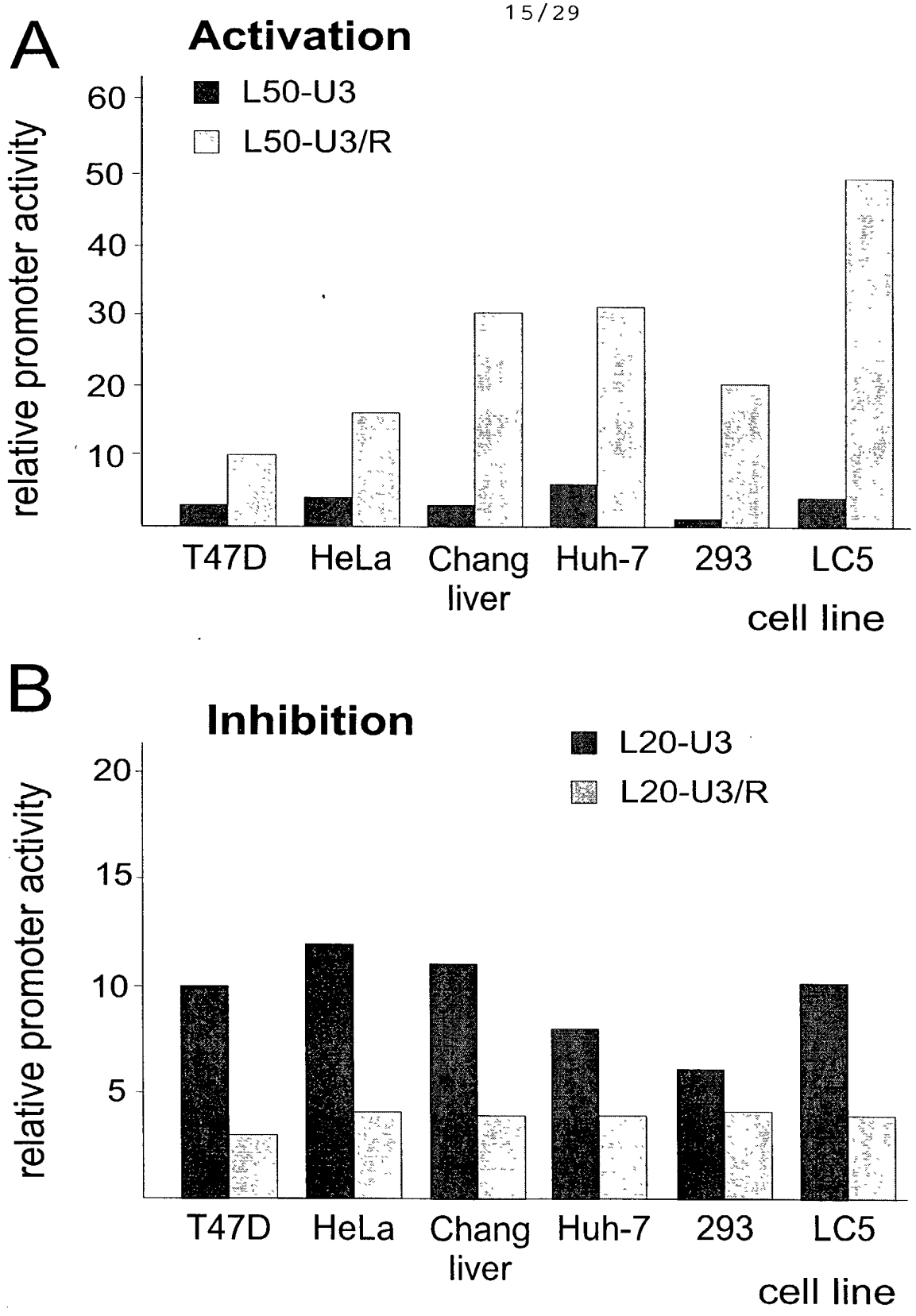


Fig. 4: LTR-R region modulates promoter activity of HERV-K-T47D related LTRs

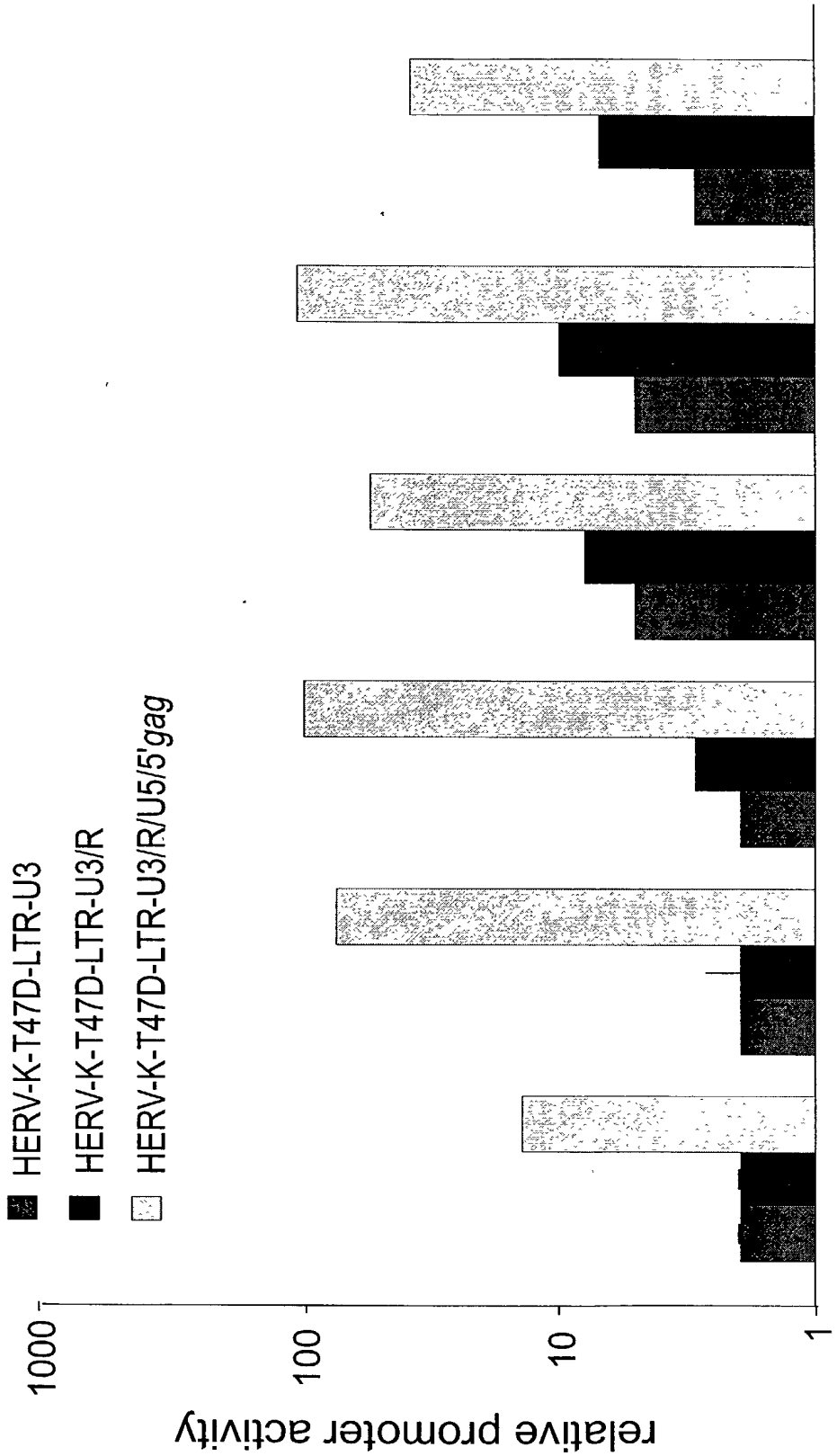


Fig. 5: Sequences downstream of LTR-R modulate promoter activity of HERV-K-T47D related LTRs

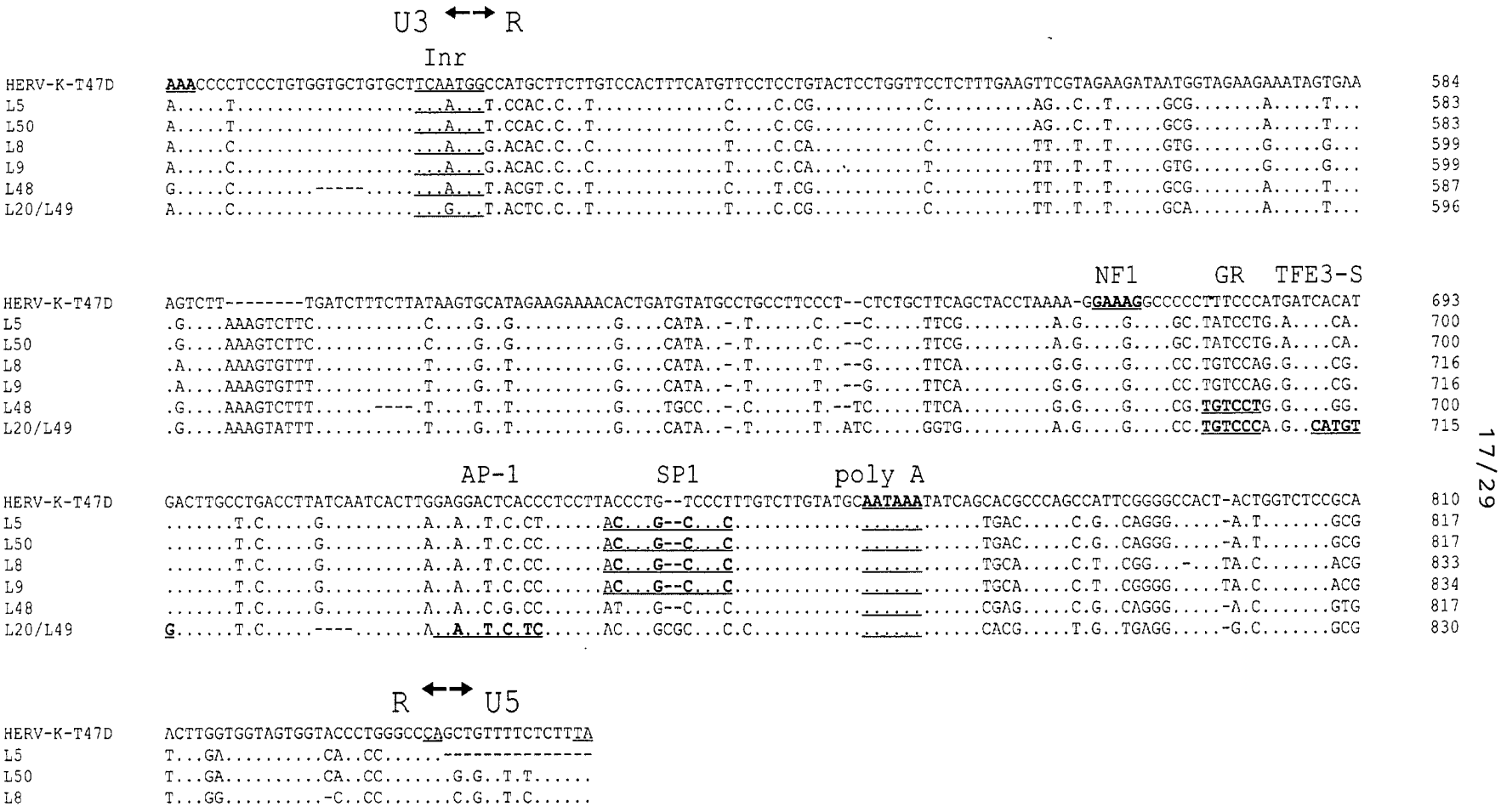
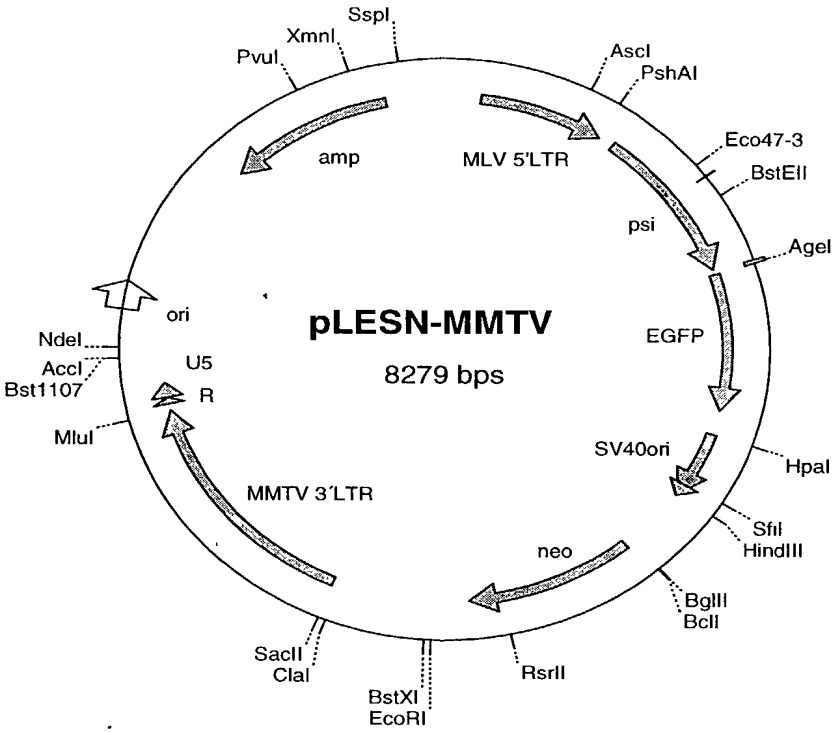


Fig. 6: Regulatory elements in the R region of HERV-K-T47D LTRs



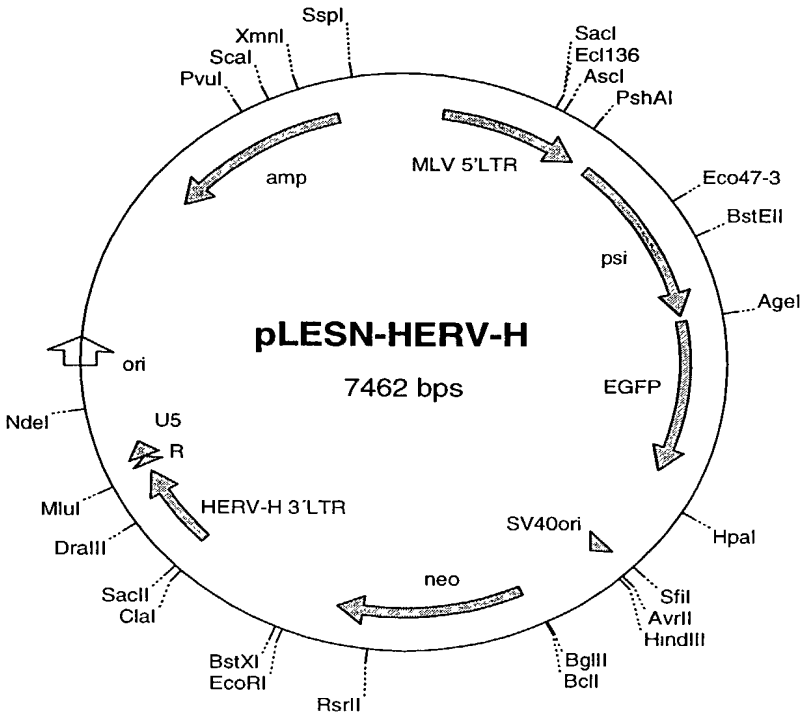


Fig.7: Retroviral ProCon vectors pLESN-MMTV and pLESN-HERV-H

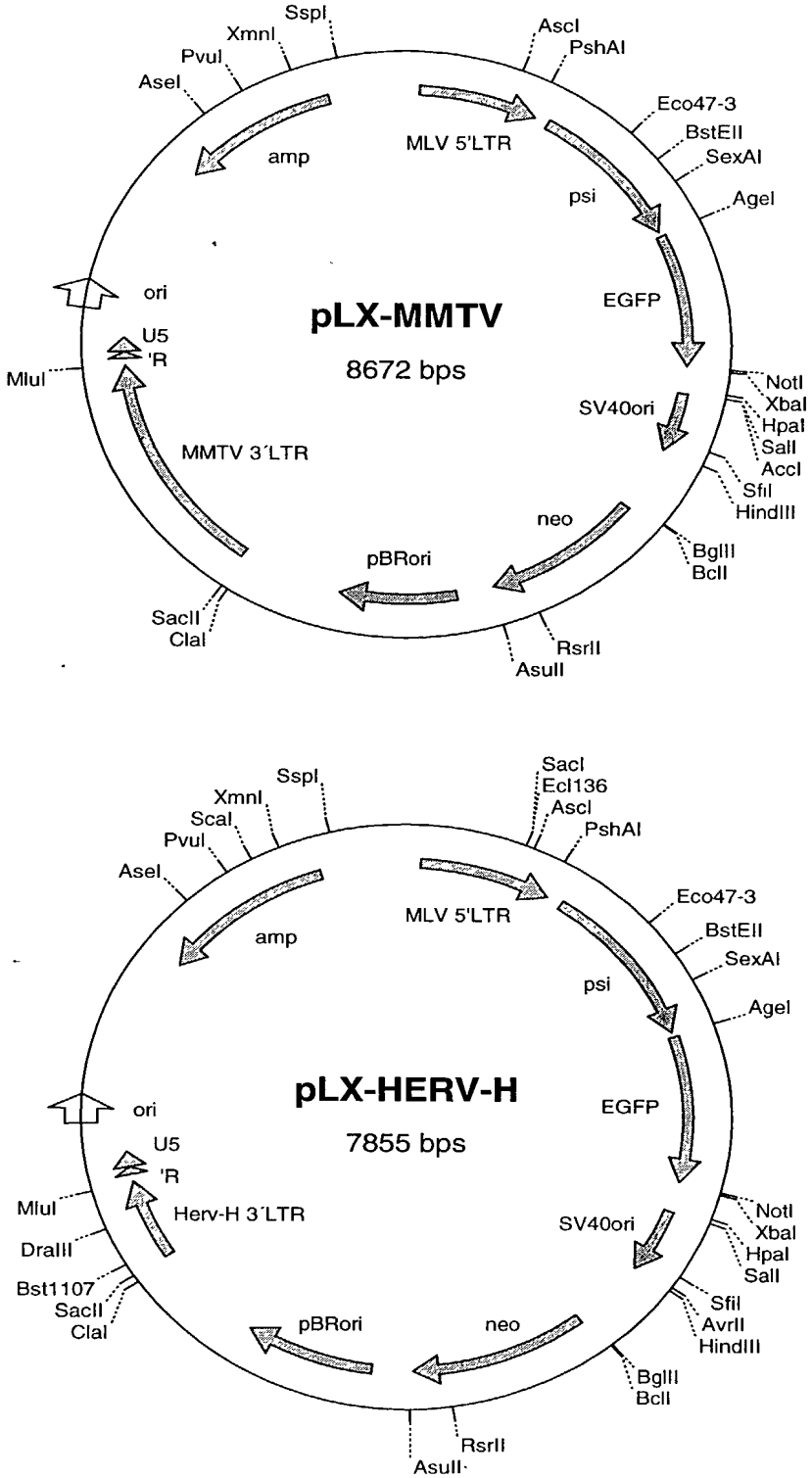
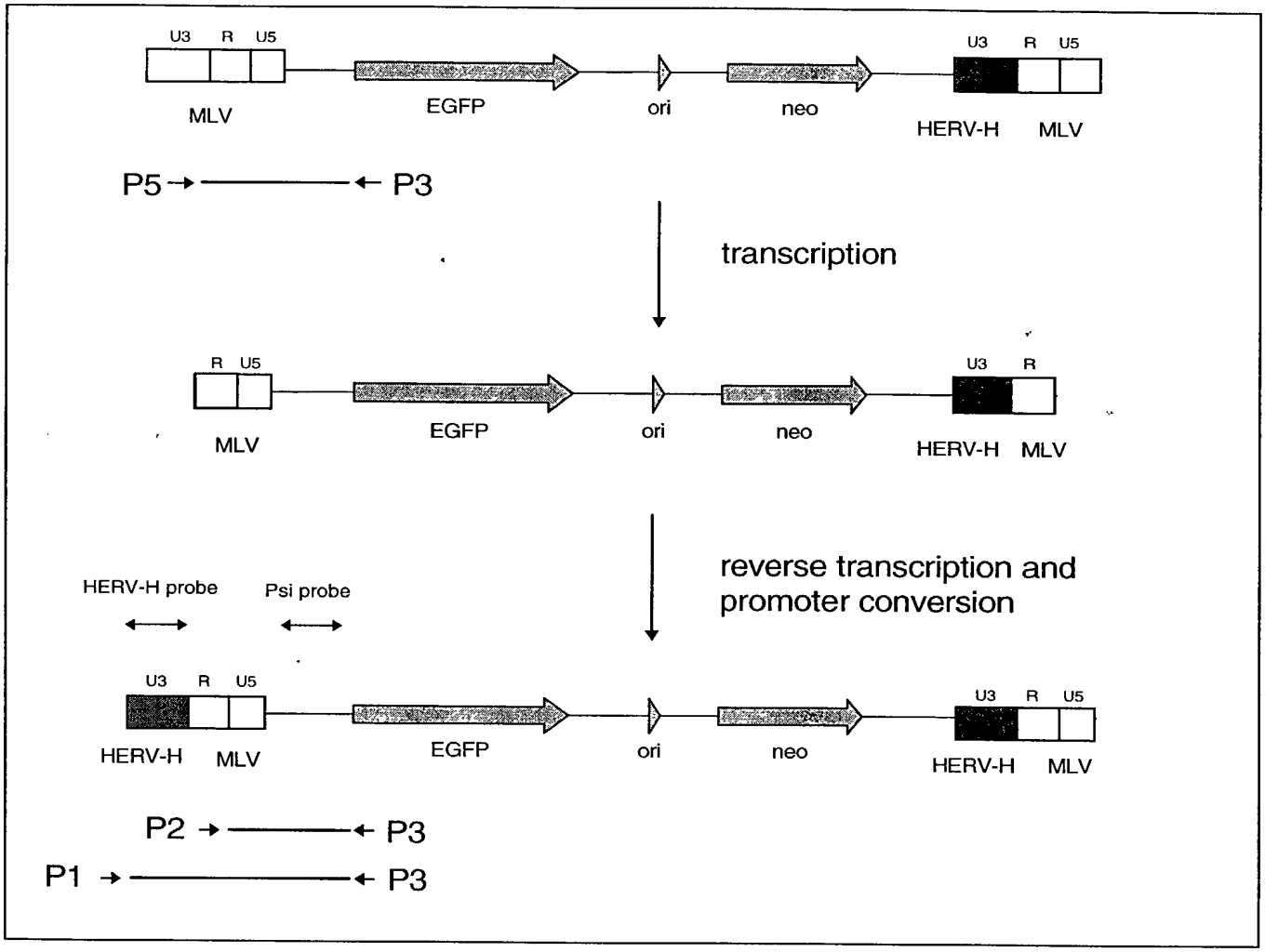


Fig.8: Retroviral ProCon vectors pLX-MMTV and pLX-HERV-H

a)



b)

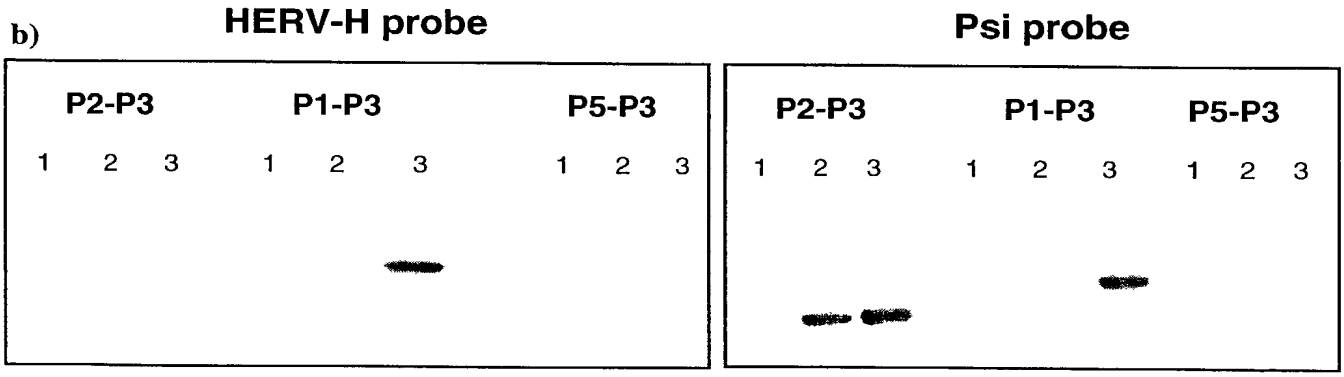
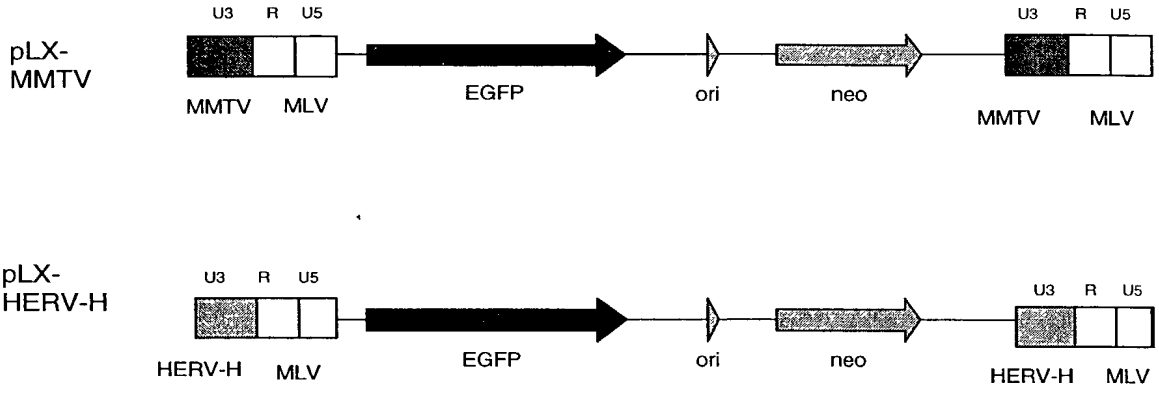


Fig. 9: a) Promoter conversion of the hybrid ProCon vectors
b) Demonstration of the correct promoter conversion with PCR and hybridization with a HERV-H and a psi probe (1:CK; 2:CK-pLX-MMTV; 3:CK-pLX-HERV-H)

a)



b)

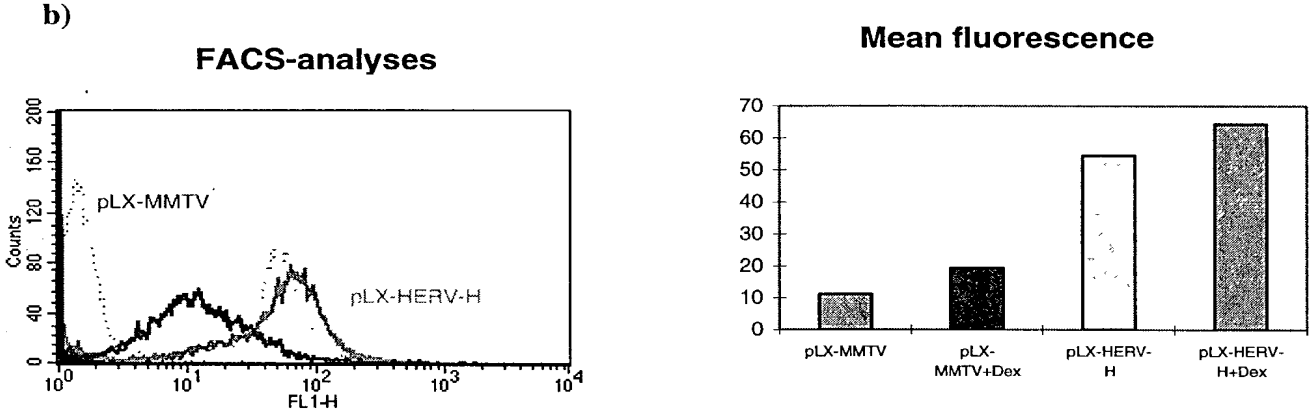


Fig.10: a) organization of the two ProCon vectors pLX-MMTV and pLX-HERV-H
b) promoter activity of the HERV-H LTR in comparison to the MMTV-LTR by infection of CrfK cells

Appendix

A. HERV-H LTR sequences

```
                1                                                    50
HERV-H L31      TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CATATCCCCT GTGACCTGCA
HERV-H HCM2     TGTCAGGCCT CTGAGCCCAA GCCAGGCCAT CGCATCCCCT GTGACTTGCA
HERV-H 19       TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CATATCCCCT GCGACCTGCA
HERV-H MP20     TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CATATCCCCT GTGACCTGCA
HERV-H CM3      TGTCAGGCCT CTGAGCCCAA GCCAAGCCAT CGCATCCCCT GTGACTTGCA
HERV-H MC16     TGTCAGGCCT CTGAGCCCAA GCC. ..... .......... ... TGCA
HERV-H CM1      TGTCAGGCCT CTGAGCCCAA GCCAAGCCAT CGCATCCCCT GTGACTTGCA
HERV-H MP23     TGTCAGGCCT CTGAGCCCAA GCCAAGCCAT CGCATCCCCT GTGACTTGCA
HERV-H H13      TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CATATCCCCA GGGACCTGCA
HERV-H H1       TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CAAATCCCCT GTGACCTGCA
HERV-H HU8      TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CATATCCC.. GTGACCTGCA
HERV-H PA7      TGTCAGGCCT CTGAGCCCAA GCTAAGCCAT CAAATCCCCT GTGACCTACA

                51                                                  100
HERV-H L31      CGTATACATC CAGATAGCC. .....TGAAG CAACTG.... ..........
HERV-H HCM2     CGTATACATC CAGATGGCC. .....TAAAG TAACTGAAGA ......TCCA
HERV-H 19       CATATACATC CAGATGGCC. .....TGAAG TAACTGAAGA ......ATCA
HERV-H MP20     CGTACACATC CAGATGGCCG GTTCCTGCCT TAACTGATGA CATTCCACCA
HERV-H CM3      CGTGTATGCC CAGATGGCC. .....TGAAG TAACTGAAGA ......ATCA
HERV-H MC16     CGTATACATC CAGATG.... .......AAG CAAGTGAAGA ......ATCA
HERV-H CM1      CGTATACGCC CAGATGGCC. .....TGAAG TAACTGAAGA ......ATCA
HERV-H MP23     CGTATACGCC CAGATGGCC. .....TGAAG TAACTGAAGA ......ATCA
HERV-H H13      CGTATACATC CAGATGGCC. .....TGAAG CAACTGAAGA ......TCCA
HERV-H H1       CGTGTACATC CAGATGACC. .....TGAAG CAACTGAAGA ......TCCA
HERV-H HU8      ..TATACATC CAGATGGCC. .....TGAAG CAACTGAAGA ......TCCA
HERV-H PA7      CGTGTACATC CAGATGACC. .....TGAAG CAACTGAAGA ......TCCA

                101                                                 150
HERV-H L31      .........T AAAAATATCC TTAACTGATG ACA.......  ..TTCCAATA
HERV-H HCM2     CAAAAGAAGT AAAAACAGCC TTAACTGATG ACA.......  ..TTCCAACA
HERV-H 19       CAAAAGAAGT GAAAATGGCC TGTTCC....  ..........  ..........
HERV-H MP20     CGAAAGAAGT GAAAATGACC TGTTCG....  ..........  ..........
HERV-H CM3      CAAAAGAAGT GAAAAGGCCC TGCCCC....  ..........  ..........
HERV-H MC16     CAAAAGAAGT GAAAATGGCC GGTTCC....  ..........  ..........
HERV-H CM1      CAAAAGAACT GAAAAGGCCC TGCCCCGCCT TAACTGATGA CATTCCACCA
HERV-H MP23     CAAAAGAACT GAAAAGGCCC TGCCCCGCCT TAACTGATGA CATTCCACCA
HERV-H H13      CAAAGGAAGT GAAAATAGCC TTAACTGATG ACA.......  ..TTCCACCA
HERV-H H1       CAAAAGAAGT GAAAGTAGCC TTAACTGATG ACA.......  ..TTCCACCA
HERV-H HU8      CAAAGGAAGT GAAAATAGCC TTAACTGATG ACA.......  ..TTCCACCA
HERV-H PA7      CAAAAGAAGT GAAAGTAGCC TTAACTGATG ACA.......  ..TTCCACCA

                151                                                 200
HERV-H L31      TTGTGATTTG TTTCTGCCCT ACCCTGACTG ATCAATGTGC TTTGTAATCT
HERV-H HCM2     TTGTGATTTG TTCCTGCCCC ACCCTAACTG ATAAATGTAC TTTGTAATCT
HERV-H 19       ..........  .........T GCCTTAACTG ATGACATTAC CTTGTGAAAT
HERV-H MP20     ..........  .........T GCCTTAACTG ATGACATTGT CTTGTGAAAT
HERV-H CM3      ..........  ..........  ACCTTAACTG AGTGATTAAC CCCATGAATT
HERV-H MC16     ..........  .........T GCCTTAACTG ATGACATTAC CTTGTGAAAT
HERV-H CM1      TGGTGATTTG TTCTTGCCCC ACCTTAACTG AGTGATTAAC CCTGTGAATT
HERV-H MP23     TGGTGATTTG TTCTTGCCCC ACCTTAACTG AGTGATTAAC CCTGTGAATT
HERV-H H13      TTGTGATTTG TTTCTGCCCC ATCCTAACTG ATCAATGTAC TTTGTAATCT
HERV-H H1       TTGTGATTTG TTCCTGCCCC ACGCTAACTG AT......AC CATATATTCT
HERV-H HU8      TTGTGATTTG TTCCTGCCCC ACGCTAACTG AT......AC CATATATTCT
HERV-H PA7      TTGTGATTTG TTCCTGCCCT ACGCTAGCTG AT......AC CATATATTCT

                201                                                 250
HERV-H L31      CCCCCACCCT TCAGAAGGCT CTTTGTAATC CTCCCCACCC TTGAGAATGG
HERV-H HCM2     CCCCCACCCT TAAGAAGGTC CTTTGTAATT CTCCCCACCC TTGAGAGTGT
HERV-H 19       TCCTTCTCCT GGCTCATCCT GGCTCAAAAG CTC..CCGCA CTGAG....C
HERV-H MP20     TCCTCCTCCT GGCTCATCCT GGCTCAAAAG CTC..CCCGA CTGAG....T
HERV-H CM3      TCCTTCCCCT GGCTCAG... .......AAG CTC..CCCCA CTGAG....C
HERV-H MC16     TCCTTCTCCT GGCTCAG... .......AAG CTC..CCCCA CTGAG....C
```

```
HERV-H CM1      TGCTTCTCCT GGCTCAG... .......AAG CTC..CCCCA CTGAG....C
HERV-H MP23     TGCTTCTCCT GGCTCAG... .......AAG CTC..CCCCA CTGAG....C
HERV-H H13      CTCCCACCCT TAAGAAGGTT CTTTGTAATT CTCCCCACCC TTGAGAGTGT
HERV-H H1       TCCCC..... .......... .......... .....CGCCC TTGAGAATGT
HERV-H HU8      TCCCC..... .......... .......... .....CGCCC TTGAGAATGT
HERV-H PA7      TCCCC..... .......... .......... .....CGCCC TTGAGAATGT

                251                                                  300
HERV-H L31      ACTTGGTGAG ATCCACCCCC TGCCTGCAAA GCATTGCCCC TAACTCCACC
HERV-H HCM2     ACTTTGTGAG ATCCACACCT GCCCACCAGA GAACAAACCC CCTTTGACTG
HERV-H 19       ACCTTGTGAC CCCTGCCTCT GCCCGCCAGA GAGCAACCCC CTCTTGACTG
HERV-H MP20     ACATTGTGAC CCCCACTCCT GCCCGCCAGA GAACAGCCCC CT.TTGACTG
HERV-H CM3      ACCTTGTGAC CCCTGCCCCT GCCCACCAGA GAACAACCCC CT.TTGACTG
HERV-H MC16     ACCTTGTGAC CCCCACTCCT CCCCGCCACA GAACAACCCC CT.TTGACTG
HERV-H CM1      ACCTTGTGAC CCCCGCCCCT GCCCACCAGA GAACAACCCC CT.TTGACTG
HERV-H MP23     ACCTTGTGAC CCCCGCCCCT GCCCACCAGA GAACAGACCC CT.TTGACTG
HERV-H H13      ACTTTGTGAG ATCCACCCCC TGCCGGCAAA ACATTGCTCC TAACCCAACC
HERV-H H1       ACTTTGTA.. .......... .......... .......... .........C
HERV-H HU8      ACTTTGTA.. .......... .......... .......... .........C
HERV-H PA7      ACTTTGTA.. .......... .......... .......... .........C

                301                                                  350
HERV-H L31      GCCTGTCCCA AAACCTATGA GAA.CTAATG ATA....... ATCCC.ACCA
HERV-H HCM2     TAATTTTCCA TTACCTTCCC TAATCCTATA AAACGGCCCC ACCCC.ATCT
HERV-H 19       TAATTTTCCT TTACCTACCT AAATCCTATA AAATGGCCCC ACTCCTATCT
HERV-H MP20     TAATTTTCCT TTATCTACCC AAATCCTATA AAACAGCCCC ACCTTTATCT
HERV-H CM3      TAATTTTCCA TTACTTTCCC AAATCCTATA AAACGGCCCC ACCCCTATCT
HERV-H MC16     TAATTTTCCA CTGCCCGCCC AAACCCTATA AAACGGTCCC ACCCC.ATCT
HERV-H CM1      TAATTTTCCA TTACCTTCCC AAATCCTATA AAACGGCCCC ACCCCTATCT
HERV-H MP23     TAATTTTCCA TTACCTTCCC AAATCCTATA AAACGGCCCC ACCCCTATCT
HERV-H H13      GCCTA.CCCC AAACCTGTAA GAA.CTAATG ATA........ATCC..ACCA
HERV-H H1       ACCTATCCC. AAACCTATAA GAA.CTAATG ATA....... ATCCT.ACCA
HERV-H HU8      ACCTATCCC. AAACCTATAA GAA.CTAATG ATA....... ATCC..ACCA
HERV-H PA7      ACCTATCCC. AAACCTATAA GAA.CTAATG ATA....... ATCCT.ACCA

                351                                                  400
HERV-H L31      CACTTTGCTG ACTCTCTTTC C...AGACTC AGCCCGGCTG CACCCAGGTG
HERV-H HCM2     CCCTTTGCTG ACTCTCTTTT C...GGACTC AGCCCGCCTG CACCCAGGTG
HERV-H 19       CCCTTCGCTG ACTCTCTTTT C...GGACTC AGCCCGCCTG TACCCAGGTG
HERV-H MP20     CCCTTGGCTG ACTCTCTTTT C...GGACTC AGCCCGCCTG CACCCAGGTG
HERV-H CM3      CCCTTCGCTG ACTCTCTTTT C...GGACTC AGCCCGCCTG CACCCAGGTG
HERV-H MC16     CCCTTCCCTG ACTCTCTTTT CTTCGGACTC AGCCCGCCTG CACCCAGGTG
HERV-H CM1      CCCTTCGCTG ACTCTCTTTT C...GGACTC AGCCCGCCTG CCCCCAGGTG
HERV-H MP23     CCCTTCGCTG ACTCTCTTTT C...GGACTC AGCCCGCCTG CCCCCAGGTG
HERV-H H13      CCCTTTGCTG ACTC..TTTT C...AGAATC AGCCCGCCTG CACCCAGGTG
HERV-H H1       CCCTTTGCTG ACTCTCTTTT T...GGACTC AGCCCGCCTG CACCCAGGTG
HERV-H HU8      CCCTTTGCTG ACTCTCTTTT T...GGACTC AGCCCGCCTG CACCCAGGTG
HERV-H PA7      CCCTTTGCTG ACTCTCTTTT T...GGACTC AGCCCGCCTG CACCCAGGTG

                401                  425
HERV-H L31      AAATAAACAG CCATGTTGCT CACAT
HERV-H HCM2     AAATAAACAG CCATGTTGCT CACAT
HERV-H 19       AAATAAACAG CCATGTTGCT CACAT
HERV-H MP20     AAATAAACAG CCATGTTGCT CACAT
HERV-H CM3      AAATAAACAG CCATGTTGCT CACAT
HERV-H MC16     AAATAAACAG CCATGTTGCT CACAT
HERV-H CM1      AAATAAACAG CCATGTTGCT CACAT
HERV-H MP23     AAATAAACAG CCATGTTGCT CACAT
HERV-H H13      AAATAAACAG CCATGTTGCT CACAT
HERV-H H1       AAATAAACAG CCATGTTGCT CACAT
HERV-H HU8      AAATAAACAG CCATGTTGCT CACAT
HERV-H PA7      AAATAAACAG CCATGTTGCT CACAT
```

B. HERV-W LTR sequences

```
                1                                                   50
HERV-T47D-W2    TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCCGA
HERV-T47D-W4    TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCCGA
HERV-T47D-W5    TGTTGAGATG GGGGACTGAG AGACAGGACT ACCTGGATTT CCTAGGCCGA
HERV-W1         TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCCAA
HERV-W10        TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTC CCTAGGCCGA
HERV-W11        TGTTGAGATG GGGGACTGAG AGACAGGACT AGTTGGATTT CCTAGGCTGG
HERV-W18        TGTTGAGATG GGGGACTGAG AGACAGGACT AGTTGGATTT CCTAGGCCGG
HERV-W2         TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCCAA
HERV-W22        TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCTGA
HERV-W23        TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCTGA
HERV-W4         TGTTGAGATG GGGGACTGAG AGACAGGACT AGTTGGATTT CCTAGGCTGG
HERV-W5         TGTTGAGATG GGGGACTGAG AGACAGGACT ACCTGGATTT CCTAGGCCAA
HERV-W6         TGTTGAGATG GGGGACTGAG AGACAGGACT AGCTGGATTT CCTAGGCCAA
HERV-W8         TGTTGAGATG GGGGACTGAG AAACAGGACT AGCAGGATTT CCTAGGCCGA

                51                                                  100
HERV-T47D-W2    CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACCGCATCC ACCTTTAAAC
HERV-T47D-W4    CTAAGAATTC CTAAGCCTAG CTGGGAAGGT GACCGCATCC ATCTTTAAAC
HERV-T47D-W5    CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACCACATCC ACCTTTAAAC
HERV-W1         CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACTACACCC ACCTTTAAAC
HERV-W10        CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACCACATCC ACCTTTAAAC
HERV-W11        CTAAGAATCC CTAAGCCTAG CTGGGAAATT GACCACGTCC ACCTTTAAAC
HERV-W18        CTAAGAATCC CTAAGCCTAG CTGGGAAATT GACCACGTCC ACCTTTAAAC
HERV-W2         CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACTACACCC ACCTTTAACC
HERV-W22        CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACCGCATCC ATCTTTAAAC
HERV-W23        CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACTACACCC ACCTTTAACC
HERV-W4         CTAAGAATCC CTAAGCCTAG CTGGGAAATT GACCACGTCC ACCTTTAAAC
HERV-W5         CTAAGAATCT CTAAGCCTAG CTGGGAAGGT GACCACATCC ACCTTTAAAC
HERV-W6         CTAAGAATCC CTAAGCCTAG CTGGGAAGGT GACTACACCC ACCTTTAACC
HERV-W8         TTAAGAATCC CTAAGCCTAG ATGGGAAGTT GACCACATCC ACCTTTAAAC

                101                                                 150
HERV-T47D-W2    ACGGGGCTCG CAACTTAGCT CACACCCAAC CAATCAGGTA GTAAAGAGGG
HERV-T47D-W4    ATGGGGCTTG CAACTTAACT CATATCTGAC CAATCAGGTA GTAAAGAGAG
HERV-T47D-W5    ACAGGGCTTG CAACTTAGCT CACACTTGAC CAGTCAGGTA GTAAAGAGAG
HERV-W1         ATGGGGCTTG CAACTTAGCT CACACCCAAC CAATCAGGTA GTAAAGAGAG
HERV-W10        ACGGGGCTTG CAACTTAGCT CATACCCAAC AAATCAGGTA GTAAAGAGAG
HERV-W11        ACGGGGCTTG CAATTTAGCT CACACCCGAC CAATCAGGTA GTAAAGGGAG
HERV-W18        ACGGGGCTTG CAATTTAGCT CACACCCGAC CAATCAGGTA GTAAAGGGAG
HERV-W2         ACTAGGCTTG CAACTTAGCT CACACCCGAC CAATCAGGTA GTAAAGAGAG
HERV-W22        ATGGGGCTTG CAACTTAACT CATATCTGAC CAATCAGGTA GTAAAGAGAG
HERV-W23        ACTAGGCTTG CAACTTAGCT CACACCCGAC CAATCAGGTA GTAAAGAGAG
HERV-W4         ACGGGGCTTG CAATTTAGCT CACACCCGAC CAATCAGGTA GTAAAGGGAG
HERV-W5         ACAGGGCTTG CAACTTAGCT CACACCCGAC CCATCAGGTA AGAAAGAGAG
HERV-W6         ACTAGGCTTG CAACTTAGCT CACACCCGAC CAATCAGGTA GTAAAGAGAG
HERV-W8         ACGGGGCTTG CAACTCAGCT CACACCCGAC CCATCAGGTA AGAAAGAGAG

                151                                                 200
HERV-T47D-W2    CTCACTAAAA TGCTAATTAG GCAAAGACAG GAGGTAAAGA AATAGCCAAT
HERV-T47D-W4    CTCACTAAAA TGCTAATTAG GCTAAAACAG GAGGCAAAGA AGTAGCCAAT
HERV-T47D-W5    CTCACTAAAA TGCTAATTAG GCTAAAACAG GAGGTAAAGA AATAGACAAT
HERV-W1         CTTGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAAAGA AATAGCCAGT
HERV-W10        CTCACTAAAA TACTGATTAG GCGAAAACAG GAGGTAAGGA AACAGCCAGT
HERV-W11        CTCACTAAAA TGCTAATTAG GGAAAAACAG GAGGTAAAGA AGTAGCCAAT
HERV-W18        CTCACTAAAA TGCTAATTAG GGAAAAACAG GAGGTAAAGA AGTAGCCAAT
HERV-W2         CTTGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAGAGA AATAGCCAAT
HERV-W22        CTTGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAAAGA AATAGCCAGT
HERV-W23        CTTGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAAAGA AATAGCCAGT
HERV-W4         CTCACTAAAA TGCTAATTAG GGAAAAACAG GAGGTAAAGA AGTAGCCAAT
HERV-W5         CCCGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAAAGA AATAGTCAAT
HERV-W6         CTTGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAAAGA AATAGCCAGT
HERV-W8         CCCGCTAAAA TGCTAATTAG GCAAAAACAG GAGGTAAAGA AATAGCCAAT
```

```
                201                                                  250
HERV-T47D-W2    CATCTATTGC CTGAGAGCAC AGCAGGAGGG ACAATGATCG GGATATAAAC
HERV-T47D-W4    CATCTGTTGC CTGACAGCAC AGCAGGAGGG ACAATGATCG GGATATAAAC
HERV-T47D-W5    CATCTATCAC CTGAGAGCAC AGTGGGAGGG ACAATGATCG GCATATAAAC
HERV-W1         CATCTATCGC CTGACAGCAC AAGGGGCGGG ACAATGATCA GGATATAAAC
HERV-W10        CATCTATCGC CTGACAGCAC AAGGGGCGGG ACAATGATCA GGATATAAAC
HERV-W11        CATCTATCGC CTGAGAGCAC AACAGGAGGG ACAATGATCA GGATATAAAC
HERV-W18        CATCTATCGC CTGAGAGCAC AACAGGAGGG ACAATGATCA GGATATAAAC
HERV-W2         CATCTATCGC CTGAGAGCAC AGCAGGAGGG ACAATGATCC GGATATAAAC
HERV-W22        CATCTATCGC CTGACAGCAC AAGGGGCGGG ACAATGATCA GGATATAAAC
HERV-W23        CATCTATCGC CTGACAGCAC AAGGGGCGGG ACAATGATCA GGATATAAAC
HERV-W4         CATCTATCGC CTGAGAGCAC AACAGGAGGG ACAATGATCA GGATATAAAC
HERV-W5         CATCTATTGC CTGAGAGCAC AGCGGGAGGG ACAATGATCA GGATATAAAC
HERV-W6         CATCTATCGC CTGACAGCAC AAGGGGCGGG ACAATGATCA GGATATAAAC
HERV-W8         CATCTATTGC CTGAGAGCAC AGCGGGAGGG ACAATGATCA GGATATAAAC

                251                                                  300
HERV-T47D-W2    CCAAGTCTTC GAGCCGGCAA TGGCTACCTT CTTTGGGTCC CCTCCCTTTG
HERV-T47D-W4    CCAGGCATTC GAGCCAGCTA CAGCTACCCT CTTTGGGTCC CCTCCCTTTG
HERV-T47D-W5    CCAGGCATTC GAGCCAGCAA CAGCAACCCC CTTTGGG... ..........
HERV-W1         TCAGGCATTC AAGCCAGCAA TGGCTACCCA CTTTGGGTCC CCTCCCATTT
HERV-W10        TCAGGCATTC AAGCCAGCAA TGGCTACCCA CTTTGGGTCC CCTCCCATTT
HERV-W11        CCAGGCATTC AAGCCAGCGG TGGCTACCCT CTTTGGGTCC CCTCCCTTTG
HERV-W18        CCAGGCATTC AAGCCAGCGG TGGCTACCCT CTTTGGGTCC CCTCCCTTTG
HERV-W2         CCAAGCATTC GAGCCAGCAA TGGCTACCCT CTTTGTGTCC CCTCCCTTTG
HERV-W22        TCAGGCATTC AAGCCAGCAA TGGCTACCCA CTTTGGGTCC CCTCCCATTT
HERV-W23        TCAGGCATTC AAGCCAGCAA TGGCTACCCA CTTTGGGTCC CCTCCCATTT
HERV-W4         CCAGGCATTC AAGCCAGCGG TGGCTACCCT CTTTGGGTCC CCTCCCTTTG
HERV-W5         CCAGGCATTC GAGCCGGCAA CGACTACCCT CTTTGGGTCC CCTCCCTTTG
HERV-W6         TCAGGCATTC AAGCCAGCAA TGGCTACCCA CTTTGGGTCC CCTCCCATTT
HERV-W8         CCAGGCATTC GAGCCGGCAA CGACTACCCT CTTTGGGTCC CCTCCCTTTG

                301                                          343
HERV-T47D-W2    TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG C..
HERV-T47D-W4    TATGGGAGCT CTGTCTTCAC TCTATTAAAT CTTGCAACTG C..
HERV-T47D-W5    ......AGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG C..
HERV-W1         TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W10        TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W11        TATGGGAGCC CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W18        TATGGAAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W2         TATGGGAGCT CTATTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W22        TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W23        TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W4         TATGGAAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W5         TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W6         TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
HERV-W8         TATGGGAGCT CTGTTTTCAC TCTATTAAAT CTTGCAACTG CAA
```

C. HERV-K LTR sequences

```
                1.................................................50
HERV-K45        GCGACCGGT: GGATC:CCGG GCCCGCGG:T ACCGTCGACT :GCAGAATTC
HERV-K27        GCGACCGGT: GGATC:CCGG GCCCGCGG:T ACCGTCGACT :GCAGAATTC
HERV-K2         GCGACCGGT: GGATC:CCGG GCCCGCGG:T ACCGTCGACT :GCAGAATTC
HERV-K1         GCGACCGGT: GGATC:CCGG GCCCGCGG:T ACCGTCGACT :GCAGAATTC
HERV-K30               GTC CCACCTCCAG CCCTAAGGCG GTTTTTCCCT ATCTCAGTAG
HERV-K10                                                      AGTAG

                51................................................100
HERV-K45        ATGGAGCATA CAATCGGGTT TTATACCGAG ACATTCCATT GCCCAGGGAC
HERV-K27        ATGGAGCATA CAATCGGGTT TTATACCGAG ACATTCCATT GCCCAGGGAC
HERV-K2         ATGGAGCATA CAATCGGGTT TTATACCGAG ACATTCCATT GCCCAGGGAC
HERV-K1         ATGGAGCATA CAATCGGGTT TTATACCGAG ACATTCCATT GCCCAGGGAC
HERV-K30        ATGGAGCATA CAATCGGGTT TTATACCGAG ACATTCCATT GCCCAGGGAC
HERV-K10        ATGGAGCATA CAATCGGGTT TTATACCGAG ACATTCCATT GCCCAGGGAC

                101...............................................150
HERV-K45        AGGCAGGAGA CAGATGCCTT CCTCTTGTCT CAACTGCAAG AGGCATTCCT
HERV-K27        AGGCAGGAGA CAGATGCCTT CCTCTTGTCT CAACTGCAAG AGGCATTCCT
HERV-K2         AGGCAGGAGA CAGATGCCTT CCTCTTGTCT CAACTGCAAG AGGCATTCCT
HERV-K1         AGGCAGGAGA CAGATGCCTT CCTCTTGTCT CAACTGCAAG AGGCATTCCT
HERV-K30        AGGCAGGAGA CAGATGCCTT CCTCTTGTCT CAACTGCAAG AGGCATTCCT
HERV-K10        AGGCAGGAGA CAGATGCCTT CCTCTTGTCT CAACTGCAAG AGGCATTCCT

                151...............................................200
HERV-K45        TCCTCTTATA CTAATCCTCC TCAGCACAGA CCCTTTACGG GTGTCGGGCT
HERV-K27        TCCTCTTATA CTAATCCTCC TCAGCACAGA CCCTTTACGG GTGTCGGGCT
HERV-K2         TCCTCTTATA CTAATCCTCC TCAGCACAGA CCCTTTACGG GTGTCGGGCT
HERV-K1         TCCTCTTATA CTAATCCTCC TCAGCACAGA CCCTTTACGG GTGTCGGGCT
HERV-K30        TCCTCTTATA CTAATCCTCC TCAGCACAGA CCCTTTACGG GTGTCGGGCT
HERV-K10        TCCTCTTTTA CTAATCCTCC TCAGCACAGA CCCTTTACAG GTGTCGGGCT

                201...............................................250
HERV-K45        GGGGGACGGT CAGGTCTTTC CCTTCCCACG AGGCCATATT TCAGACTATC
HERV-K27        GGGGGACGGT CAGGTCTTTC CCTTCCCACG AGGCCATATT TCAGACTATC
HERV-K2         GGGGGATGGT CAGGTCTTTC CCTTCCCACG AGGCCATATT TCAGACTATC
HERV-K1         GGGGGACGGT CAGGTCTTTC CCTTCCCACG AGGCCATATT TCAGACTATC
HERV-K30        GGGGGACGGT CAGGTCTTTC CCTTCCCACG AGGCCATATT TCAGACTATC
HERV-K10        GGGGGACGGT CAGGTCTTTC CCTTCCCACG AGGCCATATT TCAGACTATC

                251...............................................300
HERV-K45        ACATGGGGAG AAACCTTGGA CAATACCTGG CTTTCCTAGG CAGAGGTCCC
HERV-K27        ACATGGGGAG AAACCTTGGA CAATACCTGG CTTTCCTAGG CAGAGGTCCC
HERV-K2         ACATGGGAAG AAACCTTGGA CAATACCTGG CTTTCCTAGG CAGAGGTCCC
HERV-K1         ACATGGGGAG AAACCTTGGA CAATACCTGG CTTTCCTAGG CAGAGGTCCC
HERV-K30        ACATGGGGAG AAACCTTGGA CAATACCTGG CTTTCCTAGG CAGAGGTCCC
HERV-K10        ACATGGGGAG AAACCTTGGA CAATACCTGG CTTTCCTAGG CAGAGGTCCC

                301...............................................350
HERV-K45        TGCGGCCTTC CGCAGTTTTT GTGT:CCTGG GTACTTGAGA TTAGGGAGTG
HERV-K27        TGCGGCCTTC CGCAGTTTTT GTGT:CCTGG GTACTTGAGA TTAGGGAGTG
HERV-K2         TGCGGCCTTC CGCAGTTTTT GTGT:CCTGG GTACTTGAGA TTAGGGAGTG
HERV-K1         TGCGGCCTTC CGCAGTTTTT GTGT:CCTGG GTACTTGAGA TTAGGGAGTG
HERV-K30        TGCGGCCTTC CGCAGTTTTT GTGTCC:TGG GTACTTGAGA TTAGGGAGTG
HERV-K10        TGCGGCCTTC TGCAGTTTTT GTGTCCCTGG GTACTTGAGA TTAGGGAGTG

                351...............................................400
HERV-K45        GTGATGACTC TTAAGGAGCA TGCTGCCTTC AAGCATCTGT TTAACAAAGC
HERV-K27        GTGATGACTC TTAAGGAGCA TGCTGCCTTC AAGCATCTGT TTAACAAAGC
HERV-K2         GTGATGACTC TTAAGGAGCA TGCTGCCTTC AAGCATCTGT TTAACAAAGC
HERV-K1         GTGATGACTC TTAAGGAGCA TGCTGCCTTC AAGCATCTGT TTAACAAAGC
HERV-K30        GTGATGACTC TTAAGGAGCA TGCTGCCTTC AAGCATCTGT TTAACAAGGC
HERV-K10        GTGATGACTC TTAAGGAGCA TGCTGCCTTC AAGCATCTGT TTAACAAAGC
```

```
                401.....................................................450
HERV-K45        ACATCCTGCA CCGCCCTTAA TCCATTCAAC CCTGAGTTGA CACAGCACAC
HERV-K27        ACATCCTGCA CCGCCCTTAA TCCATTCAAC CCTGAGTTGA CACAGCACAC
HERV-K2         ACATCCTGCA CTGCCCTTAA TCCATTCAAC CCTGAGTTGA CACAGCGCAC
HERV-K1         ACATCCTGCA CCGCCCTTAA TCCATTCAAC CCTGAGTTGA CACAGCACAC
HERV-K30        ACATCCTGCA CCGCCCTTAA TCCATTCAAC CCTGAGTTGA CACAGCACAC
HERV-K10        ACATCCTGCA CCGCCCTTAA TCCATTCAAC CCTGAGTTGA CACAGCACAT

                451.....................................................550
HERV-K45        GTTTCAGAGA GCACGGGGTT GGGGGTAAGG TCATAGATTA ACAGAATCTC
HERV-K27        GTTTCAGAGA GCACGGGGTT GGGGGTAAGG TCATAGATTA ACAGAATCTC
HERV-K2         GTTTCAGAGA GCACGGGGTT GGGGGTAAGG TCATAGATTA ACAGAATCTC
HERV-K1         GTTTCAGAGA GCACGGGGTT GGGGGTAAGG TCATAGATTA ACAGAATCTC
HERV-K30        GTTTCAGAGA GCACGGGGTT GGGGGTAAGG TCATAGATTA ACAGAATCTC
HERV-K10        GTTTCAGAGA GCACGGGGTT GGGGGTAAGG TCATAGATTA ACAGAATCTC

                501.....................................................550
HERV-K45        AAGGCAGAAG AATTTTTCTT AACACATAAC AAAATGGAGT CTCCCATGTC
HERV-K27        AAGGCAGAAG AATTTTTCTT AACACATAAC AAAATGGAGT CTCCCATGTC
HERV-K2         AAGGCAGAAG AATTTTTCTT AACACATAAC AAAATGGAGT CTCCCATGTC
HERV-K1         AAGGCAGAAG AATTTTTCTT AACACATAAC AAAATGGAGT CTCCCATGTC
HERV-K30        AAGGCAGAAG AATTTTTCTT AACACATAAC AAAATGGAGT CTCCCATGTC
HERV-K10        AAGGCAGAAG AATTTTTCTT AGCACATAAC AAAATGGAGT CTCCTATGTC

                551.....................................................600
HERV-K45        TACTTCTTTC TACACAGACA CAGTAACAAT CTGATCTCTC TTGCTTTTCC
HERV-K27        TACTTCTTTC TACACAGACA CAGTAACAAT CTGATCTCTC TTGCTTTTCC
HERV-K2         TACTTCTTTC TACACAGACA CAGTAACAAT CTGATCTCTC TTGCTTTTCC
HERV-K1         TACTTCTTTC TACACAGACA CAGTAACAAT CTGATCTCTC TTGCTTTTCC
HERV-K30        TACTTCTTTC TACACAGACA CAGTAACAAT CTGATCCCTC TTGCTTTTCC
HERV-K10        TACTTCTTTC TACACAGACA CAGTAACAAT TTGATCTCTC TTGCTTTTCC

                601.....................................................650
HERV-K45        CCACATTTCC CCCTTTTCTT TTCG
HERV-K27        CCACATTTCC CCCTTTTCTT TTCGA
HERV-K2         CCACATTTCC CCCTTTTCTT TTCGACAAA
HERV-K1         CCACATTTCC CCCTTTTCTT TTCGACAAAA CCGCCAT:CT CGAGATC:TG
HERV-K30        CCACATTTCC CCCTTTTCTT ATCCATCACA CTGGCGGCCG CTCGAGCATG
HERV-K10        CCACATTTCC CCCTTTTCTT TTCGACAAAA CCGCCATC

                651.....................
HERV-K1         AGT
HERV-K30        CATCTAGAGG GCCCAATTCG CCCTATAGTG
```

HERV-K-T47D-5'LTR

TGTGGGCGAAGGATTACCCAGGTGCCGAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATATAGAAAATAGCTAG
AATAAGAATAGTTATAATAAAAATTAGATATACACATGATCATGGACATTACCAATCATTACTACAAACATTGTTAATCA
TTAGCTTTTAATATTACTCTTTGTTTTATTACTAATATAACCAAGGAATAACCGGTAGCATACGGTCAGGTGCTGAAGGG
ACATTGTGAGAAGTGACCTAGAAGGCAAGAGGTGAGCCTTCTGTCACGCCTGCATAAGGACAGCTTGAGGGCTCCTTGGT
CAAGCTGTAACACCAGTGCCTGGGAAGGCACCGTTACTTAGCAGACCATGAAAGGGAGTCTCCATTCCTTGGAGGAGTCA
GGGAAACACTATGCTCCACCAGCTTCTTGTGTATCCAGCCCTGCCCACAGTCATCCAGAGGCATAAACCCCTCCCTGTGG
TGCTGTGCTTCAATGGCCATGCTTCTTGTCCACTTTCATGTTCCTCCTGTACTCCTGGTTCCTCTTTGAAGTTCGTAGAA
GATAATGGTAGAAGAAATAGTGAAAGTCTTTGATCTTTCTTATAAGTGCATAGAAGAAAACACTGATGTATGCCTGCCTT
CCCTCTCTGCTTCAGCTACCTAAAAGGAAAGGCCCCCTTTCCCATGATCACATGACTTGCCTGACCTTATCAATCACTTG
GAGGACTCACCCTCCTTACCCTGTCCCTTTGTCTTGTATGCAATAAATATCAGCACGCCCAGCCATTCGGGGCCACTACT
GGTCTCCGCAACTTGGTGGTAGTGGTACCCTGGGCCCAGCTGTTTTCTCTTTATCTCTTTTGTCTTGTGTCTTTATTTCT
TACAATCTCTCATCTCTGCACATGGGGAGAACACCGGCAAAGCCCGTAGGGCTGGACCTTACA

L48-LTR (U3-R)

TGTGGGCGGAAGAGTACCTAGGTGCCGAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATAAAGAAAATAGAATA
AGAATAGTCATAATACAAATTAGATACAGCGATGATCATGAACAATTATCCATCATTATTATAAACATTATTAATCATTA
GCTTTTAATATTACTCTGTTGCATTAATAATATAACCTAGGAATAACCGGCAGGTATAGGGTCAGGTGCTGAAGGGACAT
TGTGAGAAGTGAATAGAAGGCAAGAGGGGAGCCTTCTGTCATGCCCGCATAAGGGCCGCTTGAGGGCCCCTTGGTCAAGC
GGTAACGCCAGTGTCTGGGAAGGCACCCGTTACTGAGCAGACCGGGAAAGGGAGTCTCCTTTCCTTGGAGGAGTCAGGGA
ACGCTCTGCTCCACCAGCTTCTTGTGGGAGGCTGGATGTTACCCAGGCCTGCCTGCAGTCATCCGGAGGCCTGAACCCCT
CCCTGTGGTGCTTCAATGGTCACGTTCCTTGTCCACTTTCATGCTCCTTCCGTACTCCTGGTTCCTCTTTGAAGTTCGTA
GTAGATAGCGGTAGAAGAAATAGTGAAAGTCTTAAAGTCTTTGATCTTATAAGTTCATAGAAGAAAACGCTGATGCCTGC
CGCCTTCTCTCTCTGCTTCAGCTACCTAAGAGGGAAGGGCCCGCTGTCCTGTGATCAGGTGACTTGCTTCACCTTGTCAA
TCACTTAGAAGACTGACCCTCCTTATCCTGCCCCCTTGTCTTGTATGCAATAAATATCAGCGAGCCCAGCCGTTCAGGGC
CACTACCGGTCTCCGTGTCTTTGTGGTAGTGGTCCCCGGGCCCAGCTGTTTTCTCTTT

L5-LTR (U3-R)

TGTGGGTGGAGGATTACCCAGGTGCCAAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATAAAAAAAAATAGAATA
AGAATAGTCATAATACAAATTAGATATAGAGATGATCATGGACAATTAGCAATCACTATTAATCTTTAGCTTTTAATATT
ACTCTTTGTTGCATTACTAATATAACCTAGGAATAACCGGTGGGTATAGGGTCAGGTGCTGAAGGGACATTGTGTGAAGT
GACCTGGAAGGCAAGAGGTGAGCCCTCTGTCACGCCCACATAAGGGCCGCTTGAGGGCTCCTTGGTCAAGTGGTAACGCC
AGTGTCTGGGAATGCACCCGTTAATTAGCAGACCGCGAAAGGGAGTCTCCTTTCCTTGGAAGAGTTGGGGAACACTCTGC
TCCACCAGCTTCTTGTGGAAGGCTGGATATTATCCAGGCCTGCGCGCAGTCATCCGGAGGCTTAAACCCTTCCCTGTGGT
GCTGTGCTTCAATGGTCCCACTCCTTGTCCACTTTCATGCTCCTCCCGTACTCCTGGTTCCTCTTTGAAGAGCGCAGTAG
ATAGCGGTAGAAGAAATAGTGAAAGTCTTAAAGTCTTCGATCTTTCTTACAAGTGCAGAGAAGAAAACGCTGACATATGC
TGCCTTCCCTCTCTGCTTCGGCTACCTAAAAGGGAAGGGCCGCCTATCCTGTAATCACATGACTTGCTTCACCTTGTCAA
TCACTTAGAAGATTCACTCTCCTTACCCTGCCCCCTTGTCTTGTATGCAATAAATATCAGTGACCCCAGCCGTTCAGGGC
CACTACTGGTCTCCGCGTCTTGATGGTAGTGGTCACCCCGGCCC

L50-LTR (U3-R)

TGTGGGTGGAGGATTACCCAGGTGCCAAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATAAAAAAAAATAGAATA
AGAATAGTCATAATACAAATTAGATATAGAGATGATCATGGACAATTAGCAATCACTATTAATCTTTAGCTTTTAATATT
ACTCTTTGTTGCATTACTAATATAACCTAGGAATAACCGGTGGGTATAGGGTCAGGTGCTGAAGGGACATTGTGAGAAGT
GACCTGGAAGGCAAGAGGTGAGCCCTCTGTCACGCCCACATAAGGGCCGCTTGAGGGCTCCTTGGTCAAGTGGTAACGCC
AGTGTCTGGGAATGCACCCGTTAATTAGCAGACCGCGAAAGGGAGTCTCCTTTCCTTGGAAGAGTTGGGGAACACTCTGC
TCCACCAGCTTCTTGTGGAAGGCTGGATATTATCCAGGCCTGCGCGCAGTCATCCGGAGGCTTAAACCCTTCCCTGTGGT
GCTGTGCTTCAATGGTCCCACTCCTTGTCCACTTTCATGCTCCTCCCGTACTCCTGGTTCCTCTTTGAAGAGCGCAGTAG
ATAGCGGTAGAAGAAATAGTGAAAGTCTTAAAGTCTTCGATCTTTCTTACAAGTGCAGAGAAGAAAACGCTGACATATGC
TGCCTTCCCTCTCTGCTTCGGCTACCTAAAAGGGAAGGGCCGCCTATCCTGTAATCACATGACTTGCTTCACCTTGTCAA
TCACTTAGAAGATTCACCCTCCTTACCCTGCCCCCTTGTCTTGTATGCAATAAATATCAGTGACCCCAGCCGTTCAGGGC
CACTACTGGTCTCCGCGTCTTGATGGTAGTGGTCACCCCGGCCCAGGTGTTTTTTCTTT

L9-LTR (966 nt)

TGTGGGTGGAGGATTACCCAGGTGCCGAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATAAAGAAAATGGTTAG
AATAAGAATAGTCATAATACAAATTAGATATAGAGATGATCATGGACAATTATCAATCATTATTATAAACATTATTAATC
ATTAGCTTTTAATATTACTCTTTGTTGCATTACTAATATAACCTAGGAATAACCGGTGGGTATAGGGTCAGGTGCTGAAA
GGACATTGGGAGAAGTGACCTAGAAGGCAAGAGGTGAGTCTTCTGTCACGCCCGCATAAGGGTTGCTTGAGGGCTCCTTG
GTCAAGTGGTAACGCCGGTGTCTGGGAAGGCACCTGTTACTTAGCCGACCACGAAAGGGAGTCTCCTTTCCTTGGAGGAG
TCAGGGCGCACTCTGCTCCACCAGCTTCTTGTGGAAGGCTGGATATTATCCAGGCCTGCCCGCAGTCATCCGGAGGCCTA
AACCCCTCCCTGTGGTGCTGTGCTTCAATGGGCACACTCCTCGTCCACTTTCATGTTCCTCCCATACTCCTGGTTTCTCT
TTGAAGTTCGTAGTAGATAGTGGTAGAAGGAATAGGGAAAATCTTAAAGTGTTTGATCTTTCTTATAAGTGCATAGAAGA
AAACGCTGACATATGCTGCCTTCTCTGTCTGCTTCAGCTACCTAAGAGGGAAGGGCCCCCTGTCCAGTGATCACGTGACT
TGCTTCACCTTGTCAATCACTTAGAAGATTCACCCTCCTTACCCTGCCCCCTTGTCTTGTATGCAATAAATATCAGTGCA
CCCAGCCTTTCGGGGCCACTTACCGGTCTCCACGTCTTGGTGGTAGTGGTCCCCCGGGCCCAGCTGTTTTCTCTTTATCT
CTTTGTCTTGTGTCTTATTTATTACAATCTCTCGTCTCCGCACACAGGGAGAACACCCGCTAAGCTCCGTAGGGCTGGAC
CCTACA

L8-LTR (938 nt)

TGTGGGTGGAGGATTACCCAGGTGCCGAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATAAAGAAAATGGTTAG
AATAAGAATAGTCATAATACAAATTAGATATAGAGATGATCATGGACAATTATCAATCATTATTATAAACATTATTAATC
ATTAGCTTTTAATATTACTCTTTGTTGCATTACTAATATAACCTAGGAATAACCGGTGGGTATAGGGTCAGGTGCTGAAG
GGACATTGGGAGAAGTGACCTAGAAGGCAAGAGGTGAGTCTTCTGTCACGCCCGCATAAGGGTTGCTTGAGGGCTCCTTG
GTCAAGTGGTAACGCCGGTGTCTGGGAAGGCACCTGTTACTTAGCCGACCACGAAAGGGAGTCTCCTTTCCTTGGAGGAG
TCAGGGCACACTCTGCTCCACCAGCTTCTTGTGGAAGGCTGGATATTATCCAGGCCTGCCCGCAGTCATCCGGAGGCCTA
AACCCCTCCCTGTGGTGCTGTGCTTCAATGGGCACACTCCTCGTCCACTTTCATGTTCCTCCCATACTCCTGGTTCCTCT
TTGAAGTTCGTAGTAGATAGTGGTAGAAGGAATAGGGAAAATCTTAAAGTGTTTGATCTTTCTTATAAGTGCATAGAAGA
AAACGCTGACATATGCTGCCTTCTCTGTCTGCTTCAGCTACCTAAGAGGGAAGGGCCCCCTGTCCAGTGATCACGTGACT
TGCTTCACCTTGTCAATCACTTAGAAGATTCACCCTCCTTACCCTGCCCCCTTGTCTTGTATGCAATAAATATCAGTGCA
CCCAGCCTTTCGGKKCACTTACCGGTCTCCACGTCTTGGTGGTAGTGGTCCCCCGGCCCAGCTGTTTTCTCTTTATCTCT
TTGTCTTGTGTCTTATTTATTACAATCTCTCGTCTCCGCACACAGGGAGAACACCCGC (Abbruch 26 nt vor
Ende der LTR)

L49-LTR = L20-LTR (963 nt)

TGTGGGCGAAAGATTACCTAGGTGCCGAGGCAAGAGACTGAAGGCACAAACTGTTTCAGTATAATAAAGAAAATAGTTAA
AATAAGAATAGTTATAATACAAATTAGATATAGAGATGATCATGGACAATTATCAATCATTATTATAAACATTAATCATT
AGCTTTTAATATTACTCTTTGTTGCTTTACTAATATAACCTAGGAATAACCGGTGGGTATAGGGTCAGGTGTTGACGGGA
TATTGTGAGAAGTGACCTAGAAGGCAAGAGGTGAGCCTTCTGTCACGCCCACATAAGGGCCGCTTGAGGGCTCTTTGGTC
AAGTGGTAACGCCAGTGTCTGTGAAGGCACCTGTTACTTAGCAGACCGCGAAAGGGAGTCTCCTTTCCTTGGAGGAGTCA
GGGAACACTCTGCTCCACCAGCTTCTTGTGGAAGGCTGGATATTATCTAGGCCTGCCCGCAGTCATCTGGAGGCCTAAAC
CCCTCCCTGTGGTGCTGTGCTTCAGTGGTCACTCTCCTTGTCCACTTTCATGTTCCTCCCGTACTCCTGGTTCCTCTTTG
AAGTTCGTAGTAGATAGCAGTAGAAGAAATAGTGAAAGTCTTAAAGTATTTGATCTTTCTTATAAGTGCATAGAAGAAAA
CGCTGACATATGCTGCCTTCTCTATCTCTGCGGTGGCTACCTAAAAGGGAAGGGCCCCCTGTCCCATGATCATGTGACTT
GCTTCACCTTATCACTTAGAAGATTCATCCTCCTTACCCTGCGCCCCCTCGTCTTGTATGCAATAAATATCAGCACGCCC
AGTCGTTTGAGGCCACTGCCGGTCTCCGCGTCTTGGTGGTAGTGGTCCCCCGGGCCCAGCTATTGTCTCTTTATCTCTTT
GTCTTGTGTCTTTATTTATTACAATCTCTTGTCTCTGCACACAGGGAGAACACCTGCTAAGCCCCGTAGGACTGGACCCT
ACA

HERV-IP-T47D

TGTTCAATTCTTTGCCTTCTACTTTTAAACTTAACTTCCTCATAAAGCAACCTTTTTCAATCACCTGCTCCACTCTGACT
CATTCTGATCACCTGCTCCACCCTGACTCATTCCGATCACCTGATCCACTGTGACTCATTCCGATTACCCGCTCCACCCT
GACTCATTCTGATTCTGATTTCCTGCTCTGCCATAACCATTTTTCCCGCCAAACCACTCACCCTGTCACTCTCTTTAAAT
TAGCCAATTGGAATTAGTTTAGCCTGTGCGGTCTAACCCTAGCCAATAGGGGACTGACACAGCAGCAGGGGCCACATGTG
TCAGGAATAAGACCCCCTTCCCCTCCCTGTCCAGATGTGTGCTCACCATTGCTCCATCTGTGAGGGCACACCCTTCTATA
GAAGTAAATTGCCTTGCTGAGAAGAAAAAAAAGAACATTTTATATTCAAGTCCTATTTCTTTTGCTGCACCGAAACTTTA
TTTATAACA

#5

Attorney's Docket No.: 10737-006001
Client's Ref. No.: P13419-DrB/la

COMBINED DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled RETROVIRAL EXPRESSION VECTORS ON THE BASIS OF HERV-LONG TERMINAL REPEAT SEQUENCES, the specification of which was filed on August 31, 2001 as Application Serial No. 09/914,665 and was amended on August 31, 2001.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose all information I know to be material to patentability in accordance with Title 37, Code of Federal Regulations, §1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) of which priority is claimed:

Country	Application No.	Filing Date	Priority Claimed
WIPO	PCT/EP00/02064	March 9, 2000	[X] Yes [] No
Germany	199 10 650.9	March 10, 1999	[X] Yes [] No

I hereby appoint the following attorneys and/or agents to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

Y. Rocky Tsao, Reg. No. 34,053
Eric L. Prahl, Reg. No. 32,590
Harold H. Fox, Reg. No. 41,498
Frank R. Occhiuti, Reg. No. 35,306
John F. Hayden, Reg. No. 37,640

Address all telephone calls to Y. ROCKY TSAO at telephone number (617) 542-5070.

Address all correspondence to Y. ROCKY TSAO at:

FISH & RICHARDSON P.C.
225 Franklin Street
Boston, Massachusetts 02110-2804

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patents issued thereon.

Attorney's Docket No.: 10737-006001
Client's Ref. No.: P13419-DrB/la

Combined Declaration and Power of Attorney

1 - ∞

Full Name of Inventor: CHRISTINE LEIB-MöSCH
Inventor's Signature: _____ Date: 7.1.02
Residence Address: Nadistraße 23, D-80809 Munchen, Germany DEX
Citizenship: German
Post Office Address: NadistraJe 23, D-80809 Munchen, Germany

2 - ∞

Full Name of Inventor: ULRIKE SCHÖN
Inventor's Signature: _____ Date: 7.1.02
Residence Address: ~~Dachauer Straße 73, D-80335 Munchen~~, Germany Neumarkter Str. 86C, 81673 München DEX
Citizenship: German
Post Office Address: ~~Dachauer StraJe 73, D-80335 Munchen~~, Germany Neumarkter Str. 86C, 81673 München

3 - ∞

Full Name of Inventor: CORINNA BAUST
Inventor's Signature: Dr. Corinna Baust Date: Dec 12, 2001
Residence Address: Schloßweg 30, D-69190 Waldorf, Germany DEX
Citizenship: German
Post Office Address: SchloJweg 30, D-69190 Waldorf, Germany

4 - ∞ *

Full Name of Inventor: ROBERT MICHAEL SALLER
Inventor's Signature: _____ Date: 14.01.2002
Residence Address: Jutastraße 22, D-80636, Munchen, Germany DEX
Citizenship: German
Post Office Address: JutastraJe 22, D-80636, Munchen, Germany

SEQUENCE LISTING

<110> GSF - Forschungszentrum für Umwelt und Gesundheit GmbH

<120> Retroviral expression vectors on the basis of
HERV-LTR-sequences

<130> P12088

<140> PCT/EP00/02064
<141> 2000-03-09

<150> DE 199 10 650.9
<151> 1999-03-10

<160> 47

<170> PatentIn Ver. 2.1

<210> 1
<211> 375
<212> DNA
<213> Human endogenous retrovirus

<400> 1
tgtcaggcct ctgagcccaa gctaagccat catatcccct gtgacctgca cgtatacatc 60
cagatagcct gaagcaactg taaaaatatc cttaactgat gacattccaa tattgtgatt 120
tgtttctgcc ctaccctgac tgatcaatgt gctttgtaat ctcccccacc cttcagaagg 180
ctctttgtaa tcctccccac ccttgagaat ggacttggtg agatccaccc cctgcctgca 240
aagcattgcc cctaactcca ccgcctgtcc caaaacctat gagaactaat gataatccca 300
ccacactttg ctgactctct ttccagactc agcccggctg cacccaggtg aaataaacag 360
ccatgttgct cacat 375

<210> 2
<211> 400
<212> DNA
<213> Human endogenous retrovirus

<400> 2
tgtcaggcct ctgagcccaa gccaggccat cgcatcccct gtgacttgca cgtatacatc 60
cagatggcct aaagtaactg aagatccaca aaagaagtaa aaacagcctt aactgatgac 120
attccaacat tgtgatttgt tcctgcccca ccctaactga taaatgtact ttgtaatctc 180
ccccaccctt aagaaggtcc tttgtaattc tccccaccct tgagagtgta ctttgtgaga 240
tccacacctg cccaccagag aacaaacccc ctttgactgt aattttccat taccttccct 300
aatcctataa aacggcccca ccccatctcc ctttgctgac tctcttttcg gactcagccc 360
gcctgcaccc aggtgaaata aacagccatg ttgctcacat 400

<210> 3
<211> 361
<212> DNA
<213> Human endogenous retrovirus

<400> 3
tgtcaggcct ctgagcccaa gctaagccat catatcccct gcgacctgca catatacatc 60
cagatggcct gaagtaactg aagaatcaca aaagaagtga aaatggcctg ttcctgcctt 120
aactgatgac attaccttgt gaaattcctt ctcctggctc atcctggctc aaaagctccc 180
gcactgagca ccttgtgacc cctgcctctg cccgccagag agcaaccccc tcttgactgt 240
aattttcctt tacctaccta aatcctataa aatggcccca ctcctatctc ccttcgctga 300
ctctcttttc ggactcagcc cgcctgtacc caggtgaaat aaacagccat gttgctcaca 360
t 361

<210> 4
<211> 372
<212> DNA
<213> Human endogenous retrovirus

<400> 4
tgtcaggcct ctgagcccaa gctaagccat catatcccct gtgacctgca cgtacacatc 60
cagatggccg gttcctgcct taactgatga cattccacca cgaaagaagt gaaaatgacc 120
tgttcgtgcc ttaactgatg acattgtctt gtgaaattcc tcctcctggc tcatcctggc 180
tcaaaagctc cccgactgag tacattgtga cccccactcc tgcccgccag agaacagccc 240
cctttgactg taatttttcct ttatctaccc aaatcctata aaacagcccc acctttatct 300
cccttggctg actctctttt cggactcagc ccgcctgcac ccaggtgaaa taaacagcca 360
tgttgctcac at 372

<210> 5
<211> 349
<212> DNA
<213> Human endogenous retrovirus

<400> 5
tgtcaggcct ctgagcccaa gccaagccat cgcatcccct gtgacttgca cgtgtatgcc 60
cagatggcct gaagtaactg aagaatcaca aaagaagtga aaaggccctg ccccacctta 120
actgagtgat taaccccatg aatttccttc ccctggctca gaagctcccc cactgagcac 180
cttgtgaccc ctgcccctgc ccaccagaga acaaccccct ttgactgtaa ttttccatta 240
ctttcccaaa tcctataaaa cggccccacc cctatctccc ttcgctgact ctcttttcgg 300
actcagcccg cctgcaccca ggtgaaataa acagccatgt tgctcacat 349

<210> 6
<211> 324
<212> DNA
<213> Human endogenous retrovirus

<400> 6
tgtcaggcct ctgagcccaa gcctgcacgt atacatccag atgaagcaag tgaagaatca 60
caaaagaagt gaaaatggcc ggttcctgcc ttaactgatg acattacctt gtgaaattcc 120
ttctcctggc tcagaagctc ccccactgag caccttgtga cccccactcc tccccgccac 180
agaacaaccc cctttgactg taatttttcca ctgcccgccc aaaccctata aaacggtccc 240
accccatctc ccttccctga ctctcttttc ttcggactca gcccgcctgc acccaggtga 300
aataaacagc catgttgctc acat 324

<210> 7
<211> 393
<212> DNA
<213> Human endogenous retrovirus

<400> 7
tgtcaggcct ctgagcccaa gccaagccat cgcatcccct gtgacttgca cgtatacgcc 60
cagatggcct gaagtaactg aagaatcaca aaagaactga aaaggccctg ccccgcctta 120
actgatgaca ttccaccatg gtgatttgtt cttgccccac cttaactgag tgattaaccc 180
tgtgaatttg cttctcctgg ctcagaagct cccccactga gcaccttgtg acccccgccc 240
ctgcccacca gagaacaacc ccctttgact gtaattttcc attaccttcc caaatcctat 300
aaaacggccc cacccctatc tcccttcgct gactctcttt tcggactcag cccgcctgcc 360
cccaggtgaa ataaacagcc atgttgctca cat 393

<210> 8
<211> 393
<212> DNA
<213> Human endogenous retrovirus

<400> 8
tgtcaggcct ctgagcccaa gccaagccat cgcatcccct gtgacttgca cgtatacgcc 60
cagatggcct gaagtaactg aagaatcaca aaagaactga aaaggccctg ccccgcctta 120
actgatgaca ttccaccatg gtgatttgtt cttgccccac cttaactgag tgattaaccc 180
tgtgaatttg cttctcctgg ctcagaagct cccccactga gcaccttgtg acccccgccc 240
ctgcccacca gagaacagac ccctttgact gtaattttcc attaccttcc caaatcctat 300
aaaacggccc cacccctatc tcccttcgct gactctcttt tcggactcag cccgcctgcc 360
cccaggtgaa ataaacagcc atgttgctca cat 393

<210> 9
<211> 388
<212> DNA
<213> Human endogenous retrovirus

<400> 9
tgtcaggcct ctgagcccaa gctaagccat catatcccca gggacctgca cgtatacatc 60
cagatggcct gaagcaactg aagatccaca aaggaagtga aaatagcctt aactgatgac 120
attccaccat tgtgatttgt ttctgcccca tcctaactga tcaatgtact ttgtaatctc 180
tcccaccctt aagaaggttc tttgtaattc tccccaccct tgagagtgta ctttgtgaga 240
tccacccccc gccggcaaaa cattgctcct aacccaaccg cctaccccaa acctgtaaga 300
actaatgata atccaccacc ctttgctgac tcttttcaga atcagcccgc ctgcacccag 360
gtgaaataaa cagccatgtt gctcacat 388

<210> 10
<211> 314
<212> DNA
<213> Human endogenous retrovirus

<400> 10
tgtcaggcct ctgagcccaa gctaagccat caaatccccc gtgacctgca cgtgtacatc 60
cagatgacct gaagcaactg aagatccaca aaagaagtga aagtagcctt aactgatgac 120
attccaccat tgtgatttgt tcctgcccca cgctaactga taccatatat tcttcccccg 180
cccttgagaa tgtactttgt acacctatcc caaacctata agaactaatg ataatcctac 240
cacccttttgc tgactctctt tttggactca gcccgcctgc acccaggtga aataaacagc 300
catgttgctc acat 314

<210> 11
<211> 309
<212> DNA
<213> Human endogenous retrovirus

<400> 11
tgtcaggcct ctgagcccaa gctaagccat catatcccgt gacctgcata tacatccaga 60
tggcctgaag caactgaaga tccacaaagg aagtgaaaat agccttaact gatgacattc 120
caccattgtg atttgttcct gccccacgct aactgatacc atatattctt cccccgccct 180
tgagaatgta ctttgtacac ctatcccaaa cctataagaa ctaatgataa tccaccaccc 240
tttgctgact ctctttttgg actcagcccg cctgcaccca ggtgaaataa acagccatgt 300
tgctcacat 309

<210> 12
<211> 314
<212> DNA
<213> Human endogenous retrovirus

<400> 12
tgtcaggcct ctgagcccaa gctaagccat caaatccccc gtgacctaca cgtgtacatc 60
cagatgacct gaagcaactg aagatccaca aaagaagtga aagtagcctt aactgatgac 120
attccaccat tgtgatttgt tcctgcccta cgctagctga taccatatat tcttcccccg 180

ccctttgagaa tgtactttgt acacctatcc caaacctata agaactaatg ataatcctac 240
caccctttgc tgactctctt tttggactca gcccgcctgc acccaggtga aataaacagc 300
catgttgctc acat 314

<210> 13
<211> 341
<212> DNA
<213> Human endogenous retrovirus

<400> 13
tgttgagatg ggggactgag agacaggact agctggattt cctaggccga ctaagaatcc 60
ctaagcctag ctgggaaggt gaccgcatcc acctttaaac acggggctcg caacttagct 120
cacacccaac caatcaggta gtaaagaggg ctcactaaaa tgctaattag gcaaagacag 180
gaggtaaaga aatagccaat catctattgc ctgagagcac agcaggaggg acaatgatcg 240
ggatataaac ccaagtcttc gagccggcaa tggctacctt ctttgggtcc cctccctttg 300
tatgggagct ctgttttcac tctattaaat cttgcaactg c 341

<210> 14
<211> 341
<212> DNA
<213> Human endogenous retrovirus

<400> 14
tgttgagatg ggggactgag agacaggact agctggattt cctaggccga ctaagaattc 60
ctaagcctag ctgggaaggt gaccgcatcc atctttaaac atggggcttg caacttaact 120
catatctgac caatcaggta gtaaagagag ctcactaaaa tgctaattag gctaaaacag 180
gaggcaaaga agtagccaat catctgttgc ctgacagcac agcaggaggg acaatgatcg 240
ggatataaac ccaggcattc gagccagcta cagctaccct ctttgggtcc cctccctttg 300
tatgggagct ctgtcttcac tctattaaat cttgcaactg c 341

<210> 15
<211> 322
<212> DNA
<213> Human endogenous retrovirus

<400> 15
tgttgagatg ggggactgag agacaggact acctggattt cctaggccga ctaagaatcc 60
ctaagcctag ctgggaaggt gaccacatcc acctttaaac acagggcttg caacttagct 120
cacacttgac cagtcaggta gtaaagagag ctcactaaaa tgctaattag gctaaaacag 180
gaggtaaaga aatagacaat catctatcac ctgagagcac agtgggaggg acaatgatcg 240
gcatataaac ccaggcattc gagccagcaa cagcaacccc ctttgggagc tctgttttca 300
ctctattaaa tcttgcaact gc 322

<210> 16
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 16
tgttgagatg ggggactgag agacaggact agctggattt cctaggccaa ctaagaatcc 60
ctaagcctag ctgggaaggt gactacaccc acctttaaac atggggcttg caacttagct 120
cacacccaac caatcaggta gtaaagagag cttgctaaaa tgctaattag gcaaaaacag 180
gaggtaaaga aatagccagt catctatcgc ctgacagcac aaggggcggg acaatgatca 240
ggatataaac tcaggcattc aagccagcaa tggctaccca ctttgggtcc cctcccattt 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 17
<211> 343

<212> DNA
<213> Human endogenous retrovirus

<400> 17
tgttgagatg ggggactgag agacaggact agctggattc cctaggccga ctaagaatcc 60
ctaagcctag ctgggaaggt gaccacatcc acctttaaac acggggcttg caacttagct 120
catacccaac aaatcaggta gtaaagagag ctcactaaaa tactgattag gcgaaaacag 180
gaggtaagga aacagccagt catctatcgc ctgacagcac aaggggcggg acaatgatca 240
ggatataaac tcaggcattc aagccagcaa tggctaccca ctttgggtcc cctcccattt 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 18
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 18
tgttgagatg ggggactgag agacaggact agttggattt cctaggctgg ctaagaatcc 60
ctaagcctag ctgggaaatt gaccacgtcc acctttaaac acggggcttg caatttagct 120
cacacccgac caatcaggta gtaaagggag ctcactaaaa tgctaattag ggaaaaacag 180
gaggtaaaga agtagccaat catctatcgc ctgagagcac aacaggaggg acaatgatca 240
ggatataaac ccaggcattc aagccagcgg tggctaccct ctttgggtcc cctccctttg 300
tatgggagcc ctgttttcac tctattaaat cttgcaactg caa 343

<210> 19
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 19
tgttgagatg ggggactgag agacaggact agttggattt cctaggccgg ctaagaatcc 60
ctaagcctag ctgggaaatt gaccacgtcc acctttaaac acggggcttg caatttagct 120
cacacccgac caatcaggta gtaaagggag ctcactaaaa tgctaattag ggaaaaacag 180
gaggtaaaga agtagccaat catctatcgc ctgagagcac aacaggaggg acaatgatca 240
ggatataaac ccaggcattc aagccagcgg tggctaccct ctttgggtcc cctccctttg 300
tatggaagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 20
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 20
tgttgagatg ggggactgag agacaggact agctggattt cctaggccaa ctaagaatcc 60
ctaagcctag ctgggaaggt gactacaccc acctttaacc actaggcttg caacttagct 120
cacacccgac caatcaggta gtaaagagag cttgctaaaa tgctaattag gcaaaaacag 180
gaggtagaga aatagccaat catctatcgc ctgagagcac agcaggaggg acaatgatcc 240
ggatataaac ccaagcattc gagccagcaa tggctaccct ctttgtgtcc cctccctttg 300
tatgggagct ctattttcac tctattaaat cttgcaactg caa 343

<210> 21
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 21
tgttgagatg ggggactgag agacaggact agctggattt cctaggctga ctaagaatcc 60
ctaagcctag ctgggaaggt gaccgcatcc atctttaaac atggggcttg caacttaact 120
catatctgac caatcaggta gtaaagagag cttgctaaaa tgctaattag gcaaaaacag 180

gaggtaaaga aatagccagt catctatcgc ctgacagcac aaggggcggg acaatgatca 240
ggatataaac tcaggcattc aagccagcaa tggctaccca ctttgggtcc cctcccattt 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 22
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 22
tgttgagatg ggggactgag agacaggact agctggattt cctaggctga ctaagaatcc 60
ctaagcctag ctgggaaggt gactacaccc acctttaacc actaggcttg caacttagct 120
cacacccgac caatcaggta gtaaagagag cttgctaaaa tgctaattag gcaaaaacag 180
gaggtaaaga aatagccagt catctatcgc ctgacagcac aaggggcggg acaatgatca 240
ggatataaac tcaggcattc aagccagcaa tggctaccca ctttgggtcc cctcccattt 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 23
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 23
tgttgagatg ggggactgag agacaggact agttggattt cctaggctgg ctaagaatcc 60
ctaagcctag ctgggaaatt gaccacgtcc acctttaaac acggggcttg caatttagct 120
cacacccgac caatcaggta gtaaagggag ctcactaaaa tgctaattag ggaaaaacag 180
gaggtaaaga agtagccaat catctatcgc ctgagagcac aacaggaggg acaatgatca 240
ggatataaac ccaggcattc aagccagcgg tggctaccct ctttgggtcc cctccctttg 300
tatggaagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 24
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 24
tgttgagatg ggggactgag agacaggact acctggattt cctaggccaa ctaagaatct 60
ctaagcctag ctgggaaggt gaccacatcc acctttaaac acagggcttg caacttagct 120
cacacccgac ccatcaggta agaaagagag cccgctaaaa tgctaattag gcaaaaacag 180
gaggtaaaga aatagtcaat catctattgc ctgagagcac agcgggaggg acaatgatca 240
ggatataaac ccaggcattc gagccggcaa cgactaccct ctttgggtcc cctccctttg 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 25
<211> 343
<212> DNA
<213> Human endogenous retrovirus

<400> 25
tgttgagatg ggggactgag agacaggact agctggattt cctaggccaa ctaagaatcc 60
ctaagcctag ctgggaaggt gactacaccc acctttaacc actaggcttg caacttagct 120
cacacccgac caatcaggta gtaaagagag cttgctaaaa tgctaattag gcaaaaacag 180
gaggtaaaga aatagccagt catctatcgc ctgacagcac aaggggcggg acaatgatca 240
ggatataaac tcaggcattc aagccagcaa tggctaccca ctttgggtcc cctcccattt 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 26
<211> 343

<212> DNA
<213> Human endogenous retrovirus

<400> 26
tgttgagatg ggggactgag aaacaggact agcaggattt cctaggccga ttaagaatcc 60
ctaagcctag atgggaagtt gaccacatcc acctttaaac acggggcttg caactcagct 120
cacacccgac ccatcaggta agaaagagag cccgctaaaa tgctaattag gcaaaaacag 180
gaggtaaaga aatagccaat catctattgc ctgagagcac agcgggaggg acaatgatca 240
ggatataaac ccaggcattc gagccggcaa cgactaccct ctttgggtcc cctccctttg 300
tatgggagct ctgttttcac tctattaaat cttgcaactg caa 343

<210> 27
<211> 619
<212> DNA
<213> Human endogenous retrovirus

<400> 27
gcgaccggtg gatcccgggc ccgcggtacc gtcgactgca gaattcatgg agcatacaat 60
cgggttttat accgagacat tccattgccc agggacaggc aggagacaga tgccttcctc 120
ttgtctcaac tgcaagaggc attccttcct cttatactaa tcctcctcag cacagaccct 180
ttacgggtgt cgggctgggg gacggtcagg tctttccctt cccacgaggc catatttcag 240
actatcacat ggggagaaac cttggacaat acctggcttt cctaggcaga ggtccctgcg 300
gccttccgca gtttttgtgt cctgggtact tgagattagg gagtggtgat gactcttaag 360
gagcatgctg ccttcaagca tctgtttaac aaagcacatc ctgcaccgcc cttaatccat 420
tcaaccctga gttgacacag cacacgtttc agagagcacg gggttggggg taaggtcata 480
gattaacaga atctcaaggc agaagaattt ttcttaacac ataacaaaat ggagtctccc 540
atgtctactt ctttctacac agacacagta acaatctgat ctctcttgct tttccccaca 600
tttccccctt ttcttttcg 619

<210> 28
<211> 620
<212> DNA
<213> Human endogenous retrovirus

<400> 28
gcgaccggtg gatcccgggc ccgcggtacc gtcgactgca gaattcatgg agcatacaat 60
cgggttttat accgagacat tccattgccc agggacaggc aggagacaga tgccttcctc 120
ttgtctcaac tgcaagaggc attccttcct cttatactaa tcctcctcag cacagaccct 180
ttacgggtgt cgggctgggg gacggtcagg tctttccctt cccacgaggc catatttcag 240
actatcacat ggggagaaac cttggacaat acctggcttt cctaggcaga ggtccctgcg 300
gccttccgca gtttttgtgt cctgggtact tgagattagg gagtggtgat gactcttaag 360
gagcatgctg ccttcaagca tctgtttaac aaagcacatc ctgcaccgcc cttaatccat 420
tcaaccctga gttgacacag cacacgtttc agagagcacg gggttggggg taaggtcata 480
gattaacaga atctcaaggc agaagaattt ttcttaacac ataacaaaat ggagtctccc 540
atgtctactt ctttctacac agacacagta acaatctgat ctctcttgct tttccccaca 600
tttccccctt ttcttttcga 620

<210> 29
<211> 624
<212> DNA
<213> Human endogenous retrovirus

<400> 29
gcgaccggtg gatcccgggc ccgcggtacc gtcgactgca gaattcatgg agcatacaat 60
cgggttttat accgagacat tccattgccc agggacaggc aggagacaga tgccttcctc 120
ttgtctcaac tgcaagaggc attccttcct cttatactaa tcctcctcag cacagaccct 180
ttacgggtgt cgggctgggg gatggtcagg tctttccctt cccacgaggc catatttcag 240
actatcacat gggaagaaac cttggacaat acctggcttt cctaggcaga ggtccctgcg 300
gccttccgca gtttttgtgt cctgggtact tgagattagg gagtggtgat gactcttaag 360
gagcatgctg ccttcaagca tctgtttaac aaagcacatc ctgcactgcc cttaatccat 420

tcaaccctga gttgacacag cgcacgtttc agagagcacg gggttggggg taaggtcata 480
gattaacaga atctcaaggc agaagaattt ttcttaacac ataacaaaat ggagtctccc 540
atgtctactt ctttctacac agacacagta acaatctgat ctctcttgct tttccccaca 600
tttccccctt ttcttttcga caaa 624

<210> 30
<211> 646
<212> DNA
<213> Human endogenous retrovirus

<400> 30
gcgaccggtg gatcccgggc ccgcggtacc gtcgactgca gaattcatgg agcatacaat 60
cgggttttat accgagacat tccattgccc agggacaggc aggagacaga tgccttcctc 120
ttgtctcaac tgcaagaggc attccttcct cttatactaa tcctcctcag cacagaccct 180
ttacgggtgt cgggctgggg gacggtcagg tctttccctt cccacgaggc catatttcag 240
actatcacat ggggagaaac cttggacaat acctggcttt cctaggcaga ggtccctgcg 300
gccttccgca gtttttgtgt cctgggtact tgagattagg gagtggtgat gactcttaag 360
gagcatgctg ccttcaagca tctgtttaac aaagcacatc ctgcaccgcc cttaatccat 420
tcaaccctga gttgacacag cacacgtttc agagagcacg gggttggggg taaggtcata 480
gattaacaga atctcaaggc agaagaattt ttcttaacac ataacaaaat ggagtctccc 540
atgtctactt ctttctacac agacacagta acaatctgat ctctcttgct tttccccaca 600
tttccccctt ttcttttcga caaaaccgcc atctcgagat ctgagt 646

<210> 31
<211> 672
<212> DNA
<213> Human endogenous retrovirus

<400> 31
gtcccacctc cagccctaag gcggtttttc cctatctcag tagatggagc atacaatcgg 60
gttttatacc gagacattcc attgcccagg gacaggcagg agacagatgc cttcctcttg 120
tctcaactgc aagaggcatt ccttcctctt atactaatcc tcctcagcac agacccttta 180
cgggtgtcgg gctgggggac ggtcaggtct ttcccttccc acgaggccat atttcagact 240
atcacatggg gagaaaccct ggacaatacc tggctttcct aggcagaggt ccctgcggcc 300
ttccgcagtt tttgtgtcct gggtacttga gattagggag tggtgatgac tcttaaggag 360
catgctgcct tcaagcatct gtttaacaag gcacatcctg caccgccctt aatccattca 420
accctgagtt gacacagcac acgtttcaga gagcacgggg ttgggggtaa ggtcatagat 480
taacagaatc tcaaggcaga agaatttttc ttaacacata acaaaatgga gtctcccatg 540
tctacttctt tctacacaga cacagtaaca atctgatccc tcttgctttt ccccacattt 600
ccccctttc ttatccatca cactggcggc cgctcgagca tgcatctaga gggcccaatt 660
cgccctatag tg 672

<210> 32
<211> 593
<212> DNA
<213> Human endogenous retrovirus

<400> 32
agtagatgga gcatacaatc gggttttata ccgagacatt ccattgccca gggacaggca 60
ggagacagat gccttcctct tgtctcaact gcaagaggca ttccttcctc ttttactaat 120
cctcctcagc acagacccttt tacaggtgtc gggctggggg acggtcaggt ctttcccttc 180
ccacgaggcc atatttcaga ctatcacatg gggagaaacc ttggacaata cctggctttc 240
ctaggcagag gtccctgcgg ccttctgcag tttttgtgtc cctgggtact tgagattagg 300
gagtggtgat gactcttaag gagcatgctg ccttcaagca tctgtttaac aaagcacatc 360
ctgcaccgcc cttaatccat tcaaccctga gttgacacag cacatgtttc agagagcacg 420
gggttggggg taaggtcata gattaacaga atctcaaggc agaagaattt ttcttagcac 480
ataacaaaat ggagtctcct atgtctactt ctttctacac agacacagta acaatttgat 540
ctctcttgct tttccccaca tttcccccttt ttcttttcga caaaaccgcc atc 593

<210> 33
<211> 943
<212> DNA
<213> Human endogenous retrovirus

<400> 33
tgtgggcgaa ggattaccca ggtgccgagg caagagactg aaggcacaaa ctgtttcagt 60
ataatataga aaatagctag aataagaata gttataataa aaattagata tacacatgat 120
catggacatt accaatcatt actacaaaca ttgttaatca ttagctttta atattactct 180
ttgttttatt actaatataa ccaaggaata accggtagca tacggtcagg tgctgaaggg 240
acattgtgag aagtgaccta gaaggcaaga ggtgagcctt ctgtcacgcc tgcataagga 300
cagcttgagg gctccttggt caagctgtaa caccagtgcc tgggaaggca ccgttactta 360
gcagaccatg aaagggagtc tccattcctt ggaggagtca gggaaacact atgctccacc 420
agcttcttgt gtatccagcc ctgcccacag tcatccagag gcataaaccc ctccctgtgg 480
tgctgtgctt caatggccat gcttcttgtc cactttcatg ttcctcctgt actcctggtt 540
cctctttgaa gttcgtagaa gataatggta gaagaaatag tgaaagtctt tgatctttct 600
tataagtgca tagaagaaaa cactgatgta tgcctgcctt ccctctctgc ttcagctacc 660
taaaaggaaa ggcccccttt cccatgatca catgacttgc ctgaccttat caatcacttg 720
gaggactcac cctccttacc ctgtcccttt gtcttgtatg caataaatat cagcacgccc 780
agccattcgg ggccactact ggtctccgca acttggtggt agtggtaccc tgggcccagc 840
tgttttctct ttatctcttt tgtcttgtgt ctttatttct tacaatctct catctctgca 900
catggggaga acaccggcaa agcccgtagg gctggacctt aca 943

<210> 34
<211> 389
<212> DNA
<213> Human endogenous retrovirus

<400> 34
aaacccctcc ctgtggtgct gtgcttcaat ggccatgctt cttgtccact ttcatgttcc 60
tcctgtactc ctggttcctc tttgaagttc gtagaagata atggtagaag aaatagtgaa 120
agtctttgat ctttcttata agtgcataga agaaaacact gatgtatgcc tgccttccct 180
ctctgcttca gctacctaaa aggaaaggcc ccctttccca tgatcacatg acttgcctga 240
ccttatcaat cacttggagg actcaccctc cttaccctgt ccctttgtct tgtatgcaat 300
aaatatcagc acgcccagcc attcggggcc actactggtc tccgcaactt ggtggtagtg 360
gtaccctggg cccagctgtt ttctcttta 389

<210> 35
<211> 858
<212> DNA
<213> Human endogenous retrovirus

<400> 35
tgtgggcgga agagtaccta ggtgccgagg caagagactg aaggcacaaa ctgtttcagt 60
ataataaaga aaatagaata agaatagtca taatacaaat tagatacagc gatgatcatg 120
aacaattatc catcattatt ataaacatta ttaatcatta gcttttaata ttactctgtt 180
gcattaataa tataacctag gaataaccgg caggtatagg gtcaggtgct gaagggacat 240
tgtgagaagt gaatagaagg caagagggga gccttctgtc atgcccgcat aagggccgct 300
tgagggcccc ttggtcaagc ggtaacgcca gtgtctggga aggcacccgt tactgagcag 360
accgggaaag ggagtctcct ttccttggag gagtcaggga acgctctgct ccaccagctt 420
cttgtgggag gctggatgtt acccaggcct gcctgcagtc atccggaggc ctgaacccct 480
ccctgtggtg cttcaatggt cacgttcctt gtccactttc atgctccttc cgtactcctg 540
gttcctcttt gaagttcgta gtagatagcg gtagaagaaa tagtgaaagt cttaaagtct 600
ttgatcttat aagttcatag aagaaaacgc tgatgcctgc cgccttctct ctctgcttca 660
gctacctaag agggaagggc ccgctgtcct gtgatcaggt gacttgcttc accttgtcaa 720
tcacttagaa gactgaccct ccttatcctg ccccttgtc ttgtatgcaa taaatatcag 780
cgagcccagc cgttcagggc cactaccggt ctccgtgtct ttgtggtagt ggtccccggg 840
cccagctgtt ttctcttt 858

<210> 36

<211> 386
<212> DNA
<213> Human endogenous retrovirus

<400> 36
gaacccctcc ctgtggtgct tcaatggtca cgttccttgt ccactttcat gctccttccg 60
tactcctggt tcctctttga agttcgtagt agatagcggt agaagaaata gtgaaagtct 120
taaagtcttt gatcttataa gttcatagaa gaaaacgctg atgcctgccg ccttctctct 180
ctgcttcagc tacctaagag ggaagggccc gctgtcctgt gatcaggtga cttgcttcac 240
cttgtcaatc acttagaaga ctgaccctcc ttatcctgcc cccttgtctt gtatgcaata 300
aatatcagcg agcccagccg ttcagggcca ctaccggtct ccgtgtcttt gtggtagtgg 360
tccccgggcc cagctgtttt ctcttt 386

<210> 37
<211> 844
<212> DNA
<213> Human endogenous retrovirus

<400> 37
tgtgggtgga ggattaccca ggtgccaagg caagagactg aaggcacaaa ctgtttcagt 60
ataataaaaa aaatagaata agaatagtca taatacaaat tagatataga gatgatcatg 120
gacaattagc aatcactatt aatctttagc ttttaatatt actctttgtt gcattactaa 180
tataacctag gaataaccgg tgggtatagg gtcaggtgct gaagggacat tgtgtgaagt 240
gacctggaag gcaagaggtg agccctctgt cacgcccaca taagggccgc ttgagggctc 300
cttggtcaag tggtaacgcc agtgtctggg aatgcacccg ttaattagca gaccgcgaaa 360
gggagtctcc tttccttgga agagttgggg aacactctgc tccaccagct tcttgtggaa 420
ggctggatat tatccaggcc tgcgcgcagt catccggagg cttaaaccct tccctgtggt 480
gctgtgcttc aatggtccca ctccttgtcc actttcatgc tcctcccgta ctcctggttc 540
ctctttgaag agcgcagtag atagcggtag aagaaatagt gaaagtctta aagtcttcga 600
tctttcttac aagtgcagag aagaaaacgc tgacatatgc tgccttccct ctctgcttcg 660
gctacctaaa agggaagggc cgcctatcct gtaatcacat gacttgcttc accttgtcaa 720
tcacttagaa gattcactct ccttaccctg ccccccttgtc ttgtatgcaa taaatatcag 780
tgaccccagc cgttcagggc cactactggt ctccgcgtct tgatggtagt ggtcaccccg 840
gccc 844

<210> 38
<211> 381
<212> DNA
<213> Human endogenous retrovirus

<400> 38
aaacccttcc ctgtggtgct gtgcttcaat ggtcccactc cttgtccact ttcatgctcc 60
tcccgtactc ctggttcctc tttgaagagc gcagtagata gcggtagaag aaatagtgaa 120
agtcttaaag tcttcgatct ttcttacaag tgcagagaag aaaacgctga catatgctgc 180
cttccctctc tgcttcggct acctaaaagg gaagggccgc ctatcctgta atcacatgac 240
ttgcttcacc ttgtcaatca cttagaagat tcactctcct taccctgccc ccttgtcttg 300
tatgcaataa atatcagtga ccccagccgt tcagggccac tactggtctc cgcgtcttga 360
tggtagtggt caccccggcc c 381

<210> 39
<211> 859
<212> DNA
<213> Human endogenous retrovirus

<400> 39
tgtgggtgga ggattaccca ggtgccaagg caagagactg aaggcacaaa ctgtttcagt 60
ataataaaaa aaatagaata agaatagtca taatacaaat tagatataga gatgatcatg 120
gacaattagc aatcactatt aatctttagc ttttaatatt actctttgtt gcattactaa 180
tataacctag gaataaccgg tgggtatagg gtcaggtgct gaagggacat tgtgagaagt 240
gacctggaag gcaagaggtg agccctctgt cacgcccaca taagggccgc ttgagggctc 300

cttggtcaag tggtaacgcc agtgtctggg aatgcacccg ttaattagca gaccgcgaaa 360
gggagtctcc tttccttgga agagttgggg aacactctgc tccaccagct tcttgtggaa 420
ggctggatat tatccaggcc tgcgcgcagt catccggagg cttaaaccct tccctgtggt 480
gctgtgcttc aatggtccca ctccttgtcc actttcatgc tcctcccgta ctcctggttc 540
ctctttgaag agcgcagtag atagcggtag aagaaatagt gaaagtctta aagtcttcga 600
tctttcttac aagtgcagag aagaaaacgc tgacatatgc tgccttccct ctctgcttcg 660
gctacctaaa agggaagggc cgcctatcct gtaatcacat gacttgcttc accttgtcaa 720
tcacttagaa gattcaccct ccttaccctg ccccttgtc ttgtatgcaa taaatatcag 780
tgaccccagc cgttcagggc cactactggt ctccgcgtct tgatggtagt ggtcaccccg 840
gcccaggtgt tttttcttt 859

<210> 40
<211> 396
<212> DNA
<213> Human endogenous retrovirus

<400> 40
aaacccttcc ctgtggtgct gtgcttcaat ggtcccactc cttgtccact ttcatgctcc 60
tcccgtactc ctggttcctc tttgaagagc gcagtagata gcggtagaag aaatagtgaa 120
agtcttaaag tcttcgatct ttcttacaag tgcagagaag aaaacgctga catatgctgc 180
cttccctctc tgcttcggct acctaaaagg gaagggccgc ctatcctgta atcacatgac 240
ttgcttcacc ttgtcaatca cttagaagat tcaccctcct taccctgccc ccttgtcttg 300
tatgcaataa atatcagtga ccccagccgt tcagggccac tactggtctc cgcgtcttga 360
tggtagtggt caccccggcc caggtgtttt ttcttt 396

<210> 41
<211> 966
<212> DNA
<213> Human endogenous retrovirus

<400> 41
tgtgggtgga ggattaccca ggtgccgagg caagagactg aaggcacaaa ctgtttcagt 60
ataataaaga aaatggttag aataagaata gtcataatac aaattagata tagagatgat 120
catggacaat tatcaatcat tattataaac attattaatc attagctttt aatattactc 180
tttgttgcat tactaatata acctaggaat aaccggtggg tatagggtca ggtgctgaaa 240
ggacattggg agaagtgacc tagaaggcaa gaggtgagtc ttctgtcacg cccgcataag 300
ggttgcttga gggctccttg gtcaagtggt aacgccggtg tctgggaagg cacctgttac 360
ttagccgacc acgaaaggga gtctcctttc cttggaggag tcagggcgca ctctgctcca 420
ccagcttctt gtggaaggct ggatattatc caggcctgcc cgcagtcatc cggaggccta 480
aacccctccc tgtggtgctg tgcttcaatg ggcacactcc tcgtccactt tcatgttcct 540
cccatactcc tggtttctct ttgaagttcg tagtagatag tggtagaagg aatagggaaa 600
atcttaaagt gtttgatctt tcttataagt gcatagaaga aaacgctgac atatgctgcc 660
ttctctgtct gcttcagcta cctaagaggg aagggccccc tgtccagtga tcacgtgact 720
tgcttcacct tgtcaatcac ttagaagatt caccctcctt accctgcccc cttgtcttgt 780
atgcaataaa tatcagtgca cccagccttt cggggccact taccggtctc cacgtcttgg 840
tggtagtggt cccccgggcc cagctgtttt ctctttatct ctttgtcttg tgtcttattt 900
attacaatct ctcgtctccg cacacaggga gaacacccgc taagctccgt agggctggac 960
cctaca 966

<210> 42
<211> 398
<212> DNA
<213> Human endogenous retrovirus

<400> 42
aaacccctcc ctgtggtgct gtgcttcaat gggcacactc ctcgtccact ttcatgttcc 60
tcccatactc ctggtttctc tttgaagttc gtagtagata gtggtagaag gaatagggaa 120
aatcttaaag tgtttgatct ttcttataag tgcatagaag aaaacgctga catatgctgc 180
cttctctgtc tgcttcagct acctaagagg gaagggcccc ctgtccagtg atcacgtgac 240
ttgcttcacc ttgtcaatca cttagaagat tcaccctcct taccctgccc ccttgtcttg 300

tatgcaataa atatcagtgc acccagcctt tcggggccac ttaccggtct ccacgtcttg 360
gtggtagtgg tccccccgggc ccagctgttt tctcttta 398

<210> 43
<211> 938
<212> DNA
<213> Human endogenous retrovirus

<400> 43
tgtgggtgga ggattaccca ggtgccgagg caagagactg aaggcacaaa ctgtttcagt 60
ataataaaga aaatggttag aataagaata gtcataatac aaattagata tagagatgat 120
catggacaat tatcaatcat tattataaac attattaatc attagctttt aatattactc 180
tttgttgcat tactaatata acctaggaat aaccggtggg tatagggtca ggtgctgaag 240
ggacattggg agaagtgacc tagaaggcaa gaggtgagtc ttctgtcacg cccgcataag 300
ggttgcttga gggctccttg gtcaagtggt aacgccggtg tctgggaagg cacctgttac 360
ttagccgacc acgaaaggga gtctcctttc cttggaggag tcagggcaca ctctgctcca 420
ccagcttctt gtggaaggct ggatattatc caggcctgcc cgcagtcatc cggaggccta 480
aacccctccc tgtggtgctg tgcttcaatg ggcacactcc tcgtccactt tcatgttcct 540
cccatactcc tggttcctct ttgaagttcg tagtagatag tggtagaagg aatagggaaa 600
atcttaaagt gtttgatctt tcttataagt gcatagaaga aaacgctgac atatgctgcc 660
ttctctgtct gcttcagcta cctaagaggg aagggccccc tgtccagtga tcacgtgact 720
tgcttcacct tgtcaatcac ttagaagatt caccctcctt accctgcccc cttgtcttgt 780
atgcaataaa tatcagtgca cccagccttt cggkkcactt accggtctcc acgtcttggt 840
ggtagtggtc ccccggccca gctgttttct ctttatctct ttgtcttgtg tcttatttat 900
tacaatctct cgtctccgca cacagggaga acacccgc 938

<210> 44
<211> 396
<212> DNA
<213> Human endogenous retrovirus

<400> 44
aaacccctcc ctgtggtgct gtgcttcaat gggcacactc ctcgtccact ttcatgttcc 60
tcccatactc ctggttcctc tttgaagttc gtagtagata gtggtagaag gaatagggaa 120
aatcttaaag tgtttgatct ttcttataag tgcatagaag aaaacgctga catatgctgc 180
cttctctgtc tgcttcagct acctaagagg gaagggcccc ctgtccagtg atcacgtgac 240
ttgcttcacc ttgtcaatca cttagaagat tcaccctcct taccctgccc ccttgtcttg 300
tatgcaataa atatcagtgc acccagcctt tcggkkcact taccggtctc cacgtcttgg 360
tggtagtggt cccccggccc agctgttttc tcttta 396

<210> 45
<211> 963
<212> DNA
<213> Human endogenous retrovirus

<400> 45
tgtgggcgaa agattaccta ggtgccgagg caagagactg aaggcacaaa ctgtttcagt 60
ataataaaga aaatagttaa aataagaata gttataatac aaattagata tagagatgat 120
catggacaat tatcaatcat tattataaac attaatcatt agcttttaat attactcttt 180
gttgctttac taatataacc taggaataac cggtgggtat agggtcaggt gttgacggga 240
tattgtgaga agtgacctag aaggcaagag gtgagccttc tgtcacgccc acataagggc 300
cgcttgaggg ctctttggtc aagtggtaac gccagtgtct gtgaaggcac ctgttactta 360
gcagaccgcg aaagggagtc tcctttcctt ggaggagtca gggaacactc tgctccacca 420
gcttcttgtg gaaggctgga tattatctag gcctgcccgc agtcatctgg aggcctaaac 480
ccctccctgt ggtgctgtgc ttcagtggtc actctccttg tccactttca tgttcctccc 540
gtactcctgg ttcctctttg aagttcgtag tagatagcag tagaagaaat agtgaaagtc 600
ttaaagtatt tgatctttct tataagtgca tagaagaaaa cgctgacata tgctgccttc 660
tctatctctg cggtggctac ctaaaaggga agggcccccc gtcccatgat catgtgactt 720
gcttcacctt atcacttaga agattcatcc tccttaccct gcgccccctc gtcttgtatg 780
caataaatat cagcacgccc agtcgtttga ggccactgcc ggtctccgcg tcttggtggt 840

agtggtcccc cgggcccagc tattgtctct ttatctcttt gtcttgtgtc tttatttatt 900
acaatctctt gtctctgcac acagggagaa cacctgctaa gccccgtagg actggaccct 960
aca 963

<210> 46
<211> 397
<212> DNA
<213> Human endogenous retrovirus

<400> 46
aaacccctcc ctgtggtgct gtgcttcagt ggtcactctc cttgtccact ttcatgttcc 60
tcccgtactc ctggttcctc tttgaagttc gtagtagata gcagtagaag aaatagtgaa 120
agtcttaaag tatttgatct ttcttataag tgcatagaag aaaacgctga catatgctgc 180
cttctctatc tctgcggtgg ctacctaaaa gggaagggcc ccctgtccca tgatcatgtg 240
acttgcttca ccttatcact tagaagattc atcctcctta ccctgcgccc cctcgtcttg 300
tatgcaataa atatcagcac gcccagtcgt ttgaggccac tgccggtctc cgcgtcttgg 360
tggtagtggt ccccccgggcc cagctattgt ctcttta 397

<210> 47
<211> 489
<212> DNA
<213> Human endogenous retrovirus

<400> 47
tgttcaattc tttgccttct acttttaaac ttaacttcct cataaagcaa ccttttttcaa 60
tcacctgctc cactctgact cattctgatc acctgctcca ccctgactca ttccgatcac 120
ctgatccact gtgactcatt ccgattaccc gctccaccct gactcattct gattctgatt 180
tcctgctctg ccataaccat ttttcccgcc aaaccactca ccctgtcact ctctttaaat 240
tagccaattg gaattagttt agcctgtgcg gtctaaccct agccaatagg ggactgacac 300
agcagcaggg gccacatgtg tcaggaataa gacccccttc ccctccctgt ccagatgtgt 360
gctcaccatt gctccatctg tgagggcaca cccttctata gaagtaaatt gccttgctga 420
gaagaaaaaa aagaacattt tatattcaag tcctatttct tttgctgcac cgaaacttta 480
tttataaca 489